Argumentation on Alternative Reaction Pathways in Organic Chemistry

Kumulative Dissertation

Zur Erlangung des Doktorgrades

der Naturwissenschaftlichen Fachbereiche

(Fachbereich 08 – Chemie und Biologie)

der Justus-Liebig-Universität Gießen

Vorgelegt von

Leonie Sabine Lieber

Gießen, 2022

Die in dieser Dissertation dokumentierten Forschungsergebnisse entstanden in der Zeit von Februar 2019 bis April 2022 am Institut für Didaktik der Chemie an der Justus-Liebig-Universität Gießen unter der wissenschaftlichen Betreuung von Frau Prof. Dr. Nicole Graulich.

The results documented in this dissertation were obtained from February 2019 to April 2022 at the Institute of Chemistry Education at the Justus-Liebig-University Giessen under the supervision of Prof. Dr. Nicole Graulich.

Ich danke meiner wissenschaftlichen Betreuerin Frau Prof. Dr. Nicole Graulich, die mich zum Teil ihrer Arbeitsgruppe gemacht hat, mir die Möglichkeit gegeben hat, mit ihrer Unterstützung zu promovieren, mir die Freiheit zur Themenfindung und zum wissenschaftlichen Arbeiten zugestanden und bei aufgekommenen Fragen stets konstruktiv mit mir diskutiert hat. Danke für drei Jahre vertrauensvoller Zusammenarbeit.

Für Susanne, Jobst, Jana und Michael

“Only by considering alternatives – by seeking to identify what is not –

can one begin to achieve any certainty about what is.”

(Kuhn, 1992, p.164)

Kuhn, D. (1992). Thinking as Argument. *Harvard Educational Review, 62 (2)*, 155-179.

Table of Contents

I Abstract

In organic chemistry, building arguments and applying concept knowledge on reaction mechanisms are essential skills, as these skills are needed whenever decisions and judgements are made, and alternatives are weighed. Applying concept knowledge when building arguments leads to the generation of a deeper understanding of the relevant content as analytical thinking is activated. Research has revealed that students experience challenges when they build arguments or use concept knowledge. However, there has been little evidence on how to support students' individual challenges in building arguments.

Therefore, to address this gap, the first step was to design a task sequence that challenges students' typical problem-solving approach of rote memorization to shift their approach to more analytical thinking, which is the basis for building well-grounded arguments. For this purpose, we designed tasks in which twenty-nine organic chemistry students could experience a cognitive conflict by reflecting on alternative reaction pathways, aiming to promote more analytical thinking. The analysis revealed that students were able to become aware of and critically reflect on their problem-solving approach.

[Lieber, L. & Graulich, N. (2020). Thinking in Alternatives – A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry. *Journal of Chemical Education*, *97*(10), 3731-3738.]

Besides fostering more analytical thinking, we investigated students' argumentation structure when judging the plausibility of alternative reaction pathways in terms of the *claim-evidence-reasoning* model. It was investigated to which extent students used evidence and reasoning in building arguments and how they justified a change in their claim by weighing the alternative reaction pathways. The results illustrate that students need support in structuring their arguments and in applying concept knowledge to build chemically sound arguments.

[Lieber, L. & Graulich, N. (2022). Investigating Students' Argumentation when Judging the Plausibility of Alternate Reaction Pathways in Organic Chemistry. *Chemistry Education Research and Practice*, *23*, 38-54.]

In order to adequately address the diagnosed challenges when building arguments (*i.e.*, structuring arguments and applying concept knowledge), an adaptive scaffold was designed that was tailored to students' individual needs. Therefore, in a first step, students received a diagnostic scaffold to support them in building arguments while their performance of structuring arguments and applying concept knowledge was analyzed. In the second part, each student received a scaffold adapted to the area in which they experienced the greatest challenges. The evaluation of the scaffold revealed that the students expressed a positive attitude towards the adaptive scaffold and stated that they had engaged more intensively with the tasks.

[Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Students' individual needs matter – a training to adaptively address students' argumentation skills in organic chemistry, *Journal of Chemical Education, 99*(7), 2754-2761.]

In addition to the promising student feedback, a quantitative analysis was conducted to examine students' performance on the diagnostic and adapted scaffold. The comparison showed that the adapted scaffolds improved students' performance in the respective areas of support (*i.e.*, structuring arguments and applying concept knowledge) and that the gap regarding to students' performance was narrowed.

[Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns. *Chemistry Education Research and Practice*, advance article, DOI: 10.1039/d2rp00016d.]

In this dissertation, it is demonstrated that building arguments on alternative reaction pathways provides new insights into diagnosing the building of arguments and the challenges students experience as a consequence thereof, and into supporting students regarding their individual challenges with an adaptive scaffold.

II Zusammenfassung

In der Organischen Chemie sind das Bilden von Argumenten und das Anwenden von Konzeptwissen in Bezug auf Reaktionsmechanismen essentielle Fähigkeiten, da diese Fähigkeiten immer dann benötigt werden, wenn Entscheidungen getroffen, Einschätzungen gemacht und Alternativen abgewogen werden. Durch das Anwenden von Konzeptwissen beim Bilden von Argumenten kann außerdem ein tieferes Verständnis für die entsprechenden Inhalte entwickelt werden, da analytisches Denken gefördert wird. Es gibt zahlreiche Belege, dass Studierende mit Herausforderungen konfrontiert sind, wenn sie Argumente bilden oder Konzeptwissen nutzen müssen. Allerdings gibt es bisher erst wenig Erkenntnisse darüber, wie man die Studierenden bei ihren individuellen Herausforderungen beim Bilden von Argumenten unterstützen kann.

Um diese Lücke zu schließen, wurde daher in einem ersten Schritt ein Aufgabendesign entwickelt, welches die typische Problemlösestrategie des Auswendiglernens der Studierenden herausfordert, um die Studierenden zu motivieren, analytisches Denken zu nutzen, was die Grundlage für das Bilden von Argumenten darstellt. Hierzu wurden Aufgaben entwickelt, bei denen 29 Studierende der Organischen Chemie durch die Reflexion über alternative Reaktionswege einen kognitiven Konflikt erfahren konnten, was analytischeres Denken provoziert. Die Analyse zeigte, dass die Studierenden in der Lage waren, sich ihrer Problemlösestrategie bewusst zu werden und diese kritisch zu reflektieren.

[Lieber, L. & Graulich, N. (2020). Thinking in Alternatives – A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry. *Journal of Chemical Education*, *97*(10), 3731-3738.]

Neben der Förderung analytischeren Denkens wurde auch die Argumentationsstruktur der Studierenden beim Abwägen der Plausibilität der alternativen Reaktionswege mithilfe des *Claim-Evidence-Reasoning*-Modells untersucht. Dabei wurde analysiert, in welchem Ausmaß Studierende Evidence und Reasoning beim Bilden der Argumente nutzten und wie sie durch das Abwägen der alternativen Reaktionswege einen Wechsel ihres Claims begründeten. Die Ergebnisse verdeutlichen, dass die Studierenden Unterstützung beim Strukturieren ihrer Argumente und bei der Anwendung von Konzeptwissen benötigen, um gut begründete Argumente bilden zu können.

[Lieber, L. & Graulich, N. (2022). Investigating Students' Argumentation when Judging the Plausibility of Alternate Reaction Pathways in Organic Chemistry. *Chemistry Education Research and Practice*, *23*, 38-54.]

Um den diagnostizierten Herausforderungen (sprich dem Strukturieren von Argumenten und der Anwendung von Konzeptwissen) angemessen begegnen zu können, wurde ein adaptiver Scaffold entwickelt, der auf die individuellen Bedürfnisse der Studierenden zugeschnitten wurde. Daher erhielten die Studierenden in einem ersten Schritt einen Diagnosescaffold, der sie beim Bilden von Argumenten unterstützte und gleichzeitig ihre Ergebnisse analysierte. Im zweiten Teil erhielten die Studierenden einen auf ihre Herausforderungen angepassten Scaffold. Die Evaluation ergab, dass die Studierenden sich positiv gegenüber dem adaptiven Scaffold äußerten und sich intensiver mit den Aufgaben beschäftigten.

[Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Students' individual needs matter – a training to adaptively address students' argumentation skills in organic chemistry, *Journal of Chemical Education, 99*(7), 2754-2761.]

Neben der vielversprechenden Rückmeldung der Studierenden wurden in einer detaillierten quantitativen Analyse die Ergebnisse der Studierenden bei der Bearbeitung beider Scaffolds untersucht. Der Vergleich zeigte, dass sich die Studierenden durch die angepassten Scaffolds in den entsprechenden Bereichen der Unterstützung (sprich dem Strukturieren der Argumente und/oder der Anwendung von Konzeptwissen) verbesserten und dass die Lücke bezogen auf die Leistungen der Studierenden verkleinert werden konnte.

[Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns. *Chemistry Education Research and Practice*, advance article, DOI: 10.1039/d2rp00016d.]

In dieser Dissertation wird deutlich, dass das Bilden von Argumenten über alternative Reaktionswege neue Erkenntnisse über die Diagnose der Bildung von Argumenten und den damit einhergehenden Herausforderungen der Studierenden und über die Förderung der Studierenden bezogen auf ihre individuellen Herausforderungen mit einem adaptiven Scaffold liefert.

III List of Publications

Papers in Peer-reviewed Journals

Lieber, L. Ibraj, K., Caspari-Gnann, I., Graulich, N. (2022), Students' individual needs matter – a training to adaptively address students' argumentation skills in organic chemistry, *Journal of Chemical Education, 99*(7), 2754-2761.

Lieber, L. Ibraj, K., Caspari-Gnann, I., Graulich, N. (2022), Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns, *Chemistry Education Research and Practice*, advance article, DOI: 10.1039/d2rp00016d.

Groos, L., Kranz, D., Lieber, L., Maaß, K., Graulich, N. (2022), Titrieren digital oder analog – Lässt sich das experimentell-praktische Verständnis von Studierenden gleichermaßen fördern?, *ChemKON*, 29, 255-260.

Lieber, L., Graulich, N. (2022), Investigating Students' Argumentation when Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry, *Chemistry Education Research and Practice*, 23, 38-53.

Lieber, L., Graulich, N. (2020), Thinking in Alternatives—A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry*, Journal of Chemical Education*, 97 (10), 3731-3738.

Papers in Non-peer-reviewed Journals

Caspari, I., Graulich, N., Lieber, L., & Rummel, L. (2018), "Zjawisko przeskoku płomienia"—Analiza opisu zjawiska przez uczniów. *Chemia w Szkole*, *3*, 19-24.

Caspari, I., Graulich, N., Lieber, L., Rummel, L. (2017), „Die Flamme geht da runter" - Prozessbeschreibungen von Lernenden analysieren, *NiU-Unterricht Chemie,* Heft 160, 28, 19-24.

Oral Presentations

37. Fortbildungs- und Vortragstagung der Fachgruppe Chemieunterricht, DiCE meets FGCU - Analog und digital: Chemieunterricht mit Potenzial, Virtuell, 23.09.-25.09.2021, Titel: Analyse von Argumentationsstrukturen über alternative Reaktionswege in der Organischen Chemie.

GDCP Jahrestagung 2021: Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen, Virtuell, 13.09.-16.09.2021, Titel: Warum ist das (un)plausibel? Analyse von Argumentationsstrukturen in der OC.

DiCE 20: Digitalisation in Chemistry Education, Virtuell, 28. Oktober 2020, Titel: Digitale Umsetzung eines didaktischen Hochschulmoduls mithilfe von Jitsi Meet.

GDCP Jahrestagung 2020: Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch?, Virtuell, 14.- 17.09.2020, Titel: Denken in Alternativen - Ein Aufgabendesign für Lernende in der OC.

ESERA Virtual Doctoral Network 2020; 29.06-03.07.2020, Titel: Investigating the potential of alternative mechanistic reactions in learning organic chemistry.

Poster Presentations

14th Conference of the European Science Education Research Association (ESERA 2021), University of Minho-Braga, Portugal, Virtuell, 30.08.-03.09.2021, Titel: Investigating Students' Argumentation when Judging the Plausibility of Alternative Mechanistic Reactions in Organic Chemistry.

X-DBER Conference 2021, 01.03.2021-03.03.2021, Virtuell, Titel: Investigating Students' Argumentation Patterns of Alternative Reaction Pathways in Organic Chemistry.

ESERA Virtual Doctoral Network 2020; 29.06-03.07.2020, Titel: Investigating the potential of alternative mechanistic reactions in learning organic chemistry.

GDCh-Wissenschaftsforum Chemie 2019: 36. Fortbildungs- und Vortragstagung der GDCh-Fachgruppe Chemieunterricht, Aachen, 16.- 19.09.2019, Titel: Diagnostizieren, Fördern, Tutorieren – DFT.

Book Chapters

Lieber, L., Graulich, N. (2022). Epistemic stances in action – Students' reasoning process while reflecting about alternative reaction pathways in organic chemistry, In: Graulich, N., Shultz, G. (Hrsg.), Student reasoning in organic chemistry, Royal Society of Chemistry, in press.

Lieber, L., Ortmann, J., Caspari, I. & Graulich, N. (2021). Diagnostizieren, Fördern, Tutorieren - Ein Seminarkonzept zur Anwendung fachdidaktischer Methoden in Kooperation mit chemischen Fachmodulen, In: Kubsch, M., Sorge, S., Arnold, J., Graulich, N. (Hrsg.), Lehrkräftebildung neu gedacht - Ein Praxishandbuch für die Lehre in den Naturwissenschaften und deren Didaktiken, Waxmann, S. 62-71.

Lieber, L., Graulich, N. (2021), Einsatz der Wärmebildkamera im naturwissenschaftlichen Unterricht - ein Te@m Workshop, In Graf, D., Graulich, N., Lengnink, K., Martinez, H., Schreiber, C. (Hrsg.), Digitale Bildung für Lehramtsstudierende, Wiesbaden Springer FachmedienVerlag, Münster, S. 189-195.

Conference Proceedings

Lieber, L., Graulich, N. (2022), Warum ist das (un)plausibel? Analyse von Argumentationsstrukturen in der OC, In: Habig, S. (Hrgs.), Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen, Virtuell, Gesellschaft für Didaktik der Chemie und Physik, digitale Jahrestagung 2021 (accepted).

Lieber, L., Graulich, N. (2021), Denken in Alternativen - Ein Aufgabendesign für Lernende in der Organischen Chemie, In: Habig, S.(Hrsg.), Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch?, Gesellschaft für Didaktik der Chemie und Physik, digitale Jahrestagung 2020, S. 66.

Lieber, L., Graulich, N. (2020), Productive representational errors - Investigating the potential of alternative mechanistic reactions in learning organic chemistry, ESERA Virtual Doctoral Network 2020 Book of Synopses, S. 245-254.

1 Review: On the traces of diagnosing and fostering argumentation in organic chemistry

1.1 Introduction

In the year 1924, Patrick already made it clear that "good" chemistry teaching derives from the method of how facts are taught to students and how the facts are made tangible for them (Patrick, 1924). Today, the focus of many traditional learning environments is still on teaching facts and concepts (Cooper, 2015; Momsen *et al.*, 2010; Stowe *et al.*, 2021). However, the questions that should always be asked when teaching chemistry are (1) what should students actually be learning and what are students able to do with that knowledge, (2) how to find out if students have also developed an understanding of the content, and (3) how to support students in building such an understanding (Cooper and Stowe, 2018). These questions are closely related and can rarely be considered separately (Cooper and Stowe, 2018). For example, the first question refers not solely to the selection of facts or concepts students should learn, but also to support, as problem-solving strategies and scientific practices need to be learned as well (Cooper, 2015; Crooks, 1988; Pickering, 2010; Sandoval *et al.*, 2019; Songer and Gotwals, 2012). Specific questions students have to answer in chemistry may include (I) why fluorine has the highest electronegativity according to PAULING or (II) whether molecule X is more stable than molecule Y in a specific reaction. In order to answer these questions, students do not only have to be able to recall scientific principles that have been taught, but have to use methodological competencies as well. However, both aforementioned scientific questions (I + II) differ fundamentally in their nature: The first question aims at an explanation, the second at an argument. These two terms are often confounded or considered to be the same (Osborne and Patterson, 2011), hence a distinction between explanation and argumentation will be made here.

Explanations aim to answer the *why* questions and to unfold the phenomenon (Ohlsson, 2002; Osborne and Patterson, 2011). In the example of the electronegativity of fluorine, an explanation consists of a subset of descriptions (Osborne and Patterson, 2011) and intends to create a sense of increased understanding of the phenomenon by answering the *why* question (Brewer *et al.*, 1998; Wilson and Keil, 1998). The fundamental difference between explanation and argument, however, is that explanations are not driven by the need to persuade the opposite party (Osborne and Patterson, 2011), but serve to build knowledge (Ford, 2008). Explanations do not need to convince the opposite primarily because the phenomenon, which is to be explained, is not doubted, but is assumed to be true within the science community (Osborne and Patterson, 2011). This makes explanations of the phenomenon less certain than the phenomenon itself (Govier, 1987). Compared to explanations, arguments consist of a claim that has to be supported with data and justifications (Toulmin, 2003). For this reason, arguments always have a particular degree of uncertainty (Osborne and Patterson, 2011). However, the justifications used to support the claim are unquestionable, thus, building the evidence for the validity of the claim (Osborne and Patterson, 2011). Furthermore, the explanation and the argument are also directly related to each other since, for example, both practices use scientific principles. Nonetheless, they differ in their epistemic function in this regard, as explanations unfold an unquestionable phenomenon whereas the argument justifies a doubtful claim (Osborne and Patterson, 2011). Arguments are also essential in validating explanations, since there often exist many explanations for the same phenomenon. However, the claim of arguments is always that explanation A is better than explanation B and not an evaluation of the explanation itself (Osborne and Patterson, 2011). Kuhn (1992) sums up the linkage and utility of the two practices as follows:

„*Only by considering alternatives – by seeking to identify what is not – can one begin to achieve any certainty about what is*." (Kuhn, 1992, p. 164)

In chemistry, the aim is to generate explanatory hypotheses, but these are always dependent on the argumentation that a hypothesis is the best possible one at the time. Only after this process of argumentation, in which the hypothesis must withstand counterarguments, an explanatory hypothesis can be accepted as an explanation (Longino, 1990). Critical engagement with hypotheses is only fostered through the formation of arguments and thus represents one of a chemist's most important skills (Novak and Treagust, 2018; Osborne and Patterson, 2011). In doing so, a deeper understanding is generated, as building arguments and counterarguments are essential components of higher order thinking according to Bloom and colleagues (Bloom *et al.*, 1956; Kuhn, 1992; Osborne and Patterson, 2011). Argumentation occurs whenever judgements must be made, alternatives are weighed, and decisions are made (Kuhn, 1992). Therefore, the questions posed at the beginning – (1) what should students learn and what should students are able to do with the knowledge – can be answered quickly regarding argumentation, since this practice is a skill that accompanies chemists in their daily lives. Building arguments can also be used as a tool to answer the second question – (2) how to find out if students have developed an understanding of the content. This is for the reason that argumentation is closely related to concept knowledge (Songer and Gotwals, 2012), as it finds application in evaluating claims and weighing evidence (Driver *et al.*, 2000). For this reason, related to the third question – (3) how to support students – it is also equally important to support students in both concept knowledge and argumentation (Bricker and Bell, 2008). Especially by linking concept knowledge to argumentation and by actively weighing alternatives in the decision-making process, it becomes evident that argumentation plays a central role in chemistry, which is why it has been of great interest in chemistry education research recently and for decades (*e.g.*, Deng and Flynn, 2021; Hosbein *et al.*, 2021; Jiménez-Aleixandre and Erduran, 2007; Petritis *et al.*, 2021; Tüzün *et al.*, 2021; Walker *et al.*, 2019).

In order to teach students how to build arguments and to support them in this learning process, a distinction from explanation as well as general characteristics of argumentation have already been presented. In the next step, the structure of an argument will be discussed in more detail.

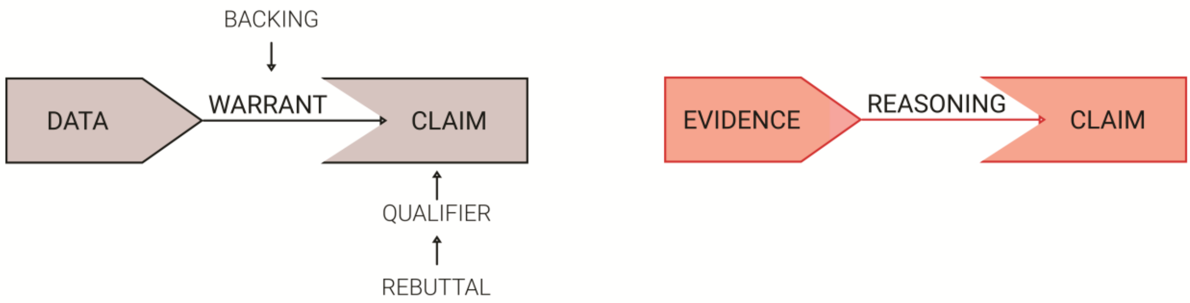


Fig. 1: Illustration of the most frequently used argumentation models in chemistry education research. On the left side is the original argumentation model from Stephen Toulmin (Toulmin, 2003), on the right side is the *claim-evidence-reasoning* model, which is a simplified version of Toulmin's argumentation model.

One of the most frequently used models of argumentation derived from Stephen Toulmin in 1958, who stated that an argument consists of six elements (see Figure 1, left side): a *claim*, which is supported by *data* based on scientific principles; a *warrant*, which reflects the justification of the *data* and acts as a bridge between *claim* and *data*; and a *backing* for the *warrant*, which is based on general norms and ethical principles. Additionally, an argument also consists of a *qualifier*, which weakens the argument and asks about the necessity of the *claim*, and a *rebuttal*, which can also be understood as a counterargument, since it represents the exceptions of the *claim* (Toulmin, 2003). This argumentation model has already been used by numerous chemistry education researchers, although simplifications have been applied (Becker *et al.*, 2013; Cruz-Ramirez de Arellano and Towns, 2014; Tüzün *et al.*, 2021). This is mainly because Toulmin's model is too complex to be used to introduce argumentation to students, which is why Toulmin already expressed that a core argument consists of only three components (*claim*, *data*, *warrant*) (Toulmin, 2003). Specific challenges in using Toulmin's argumentation model arose with students, for example, from the distinction between *data*, *warrant*, and *backing* (Erduran, 2007), and even teachers reported that the argumentation model triggered ambiguity (Lazarou and Erduran, 2020). For this reason, the *claim-evidence-reasoning* model (CER model) is currently most frequently used, as it is a simplification of Toulmin's argumentation model and consists only of the components *claim*, *evidence*, and *reasoning* (see Figure 1, right side) (McNeill and Krajcik, 2012; McNeill *et al.*, 2006). In this model, the *claim* is a problem's statement and is always in doubt, which is why the *claim* needs a justification based on scientific principles (McNeill and Krajcik, 2012; McNeill *et al.*, 2006). Here the justification is divided into two parts. The first part is the *evidence*, which can be based on multiple principles, whereby a *claim* can be supported with several pieces of *evidence* (McNeill and Krajcik, 2012). The second part of the justification is *reasoning*, which bridges *claim* and *evidence* by using scientific principles to answer the questions why the *evidence* fits the *claim* (McNeill *et al.*, 2006). Only by using all three components, a complete argument emerges, since a *claim* must be justified by *evidence* and *reasoning* (McNeill and Krajcik, 2012).

Nevertheless, the way students build arguments differs from the way teachers would expect students to build arguments. One of the greatest challenges students experience is building a complete argument, *i.e.*, justifying the claim with evidence and reasoning (Deng and Flynn, 2021; Petritis *et al.*, 2021; Stanford *et al.*, 2016). In the justification, students face the challenge of using scientific principles appropriately (McNeill and Krajcik, 2012; Walker *et al.*, 2019), leading to uncertainties about what counts as evidence (Sadler, 2004). To overcome these uncertainties, students often rely on personal feelings (Hogan and Maglienti, 2001), which results in the criteria of an argument no longer being met and the argument not resisting discussion. One reason for these challenges could be that students have difficulty activating and applying concept knowledge to build chemically sound arguments or even lack the necessary prior knowledge (Cruz-Ramirez de Arellano and Towns, 2014; Deng and Flynn, 2021; Moon *et al.*, 2016; Pabuccu and Erduran, 2017). In summary, the question of whether students have developed an understanding of the content they are learning has been diagnosed in a variety of ways. Here, it stands out that the challenges students experience in building arguments are either caused by incomplete knowledge of the structure of an argument or complete activation and application of concept knowledge, or both reasons. Due to the listed challenges of students, several support approaches for building arguments have already been developed (Henderson and Osborne, 2019; Hosbein *et al.*, 2021; Petritis *et al.*, 2021; Petritis *et al.*, 2022; Walker *et al.*, 2012). Most prominent among these approaches is scaffolding because, on the one hand, it offers many different ways of providing support, for example in the form of modelling or targeted prompting (Kang *et al.*, 2014; Van de Pol *et al.*, 2010). On the other hand, the core of scaffolding is that students are guided to solve tasks they would have failed at without help (Kang *et al.*, 2014; Lajoie, 2005; Pea, 2004; Wood *et al.*, 1976), which might be because the decision-making process is slowed down (Caspari and Graulich, 2019). This may raise the question of why further research should be investigating diagnosing and fostering argumentation when findings on both points already exist.

As mentioned before, argumentation is closely linked to the application of concept knowledge, which highlights that one should consider both aspects together. Thus, students' prior knowledge also has a major impact on the way students build arguments (de Lima Tavares *et al.*, 2010; Faize *et al.*, 2017; von Aufschnaiter *et al.*, 2008). However, both, the challenges students experience when building arguments and the prior knowledge students bring into the classroom are individual. For this reason, there is the need for students to receive support that is adapted to their individual strengths and challenges (Chen, 2014; Van de Pol *et al.*, 2010; Wood *et al.*, 1976). This need is reinforced by several researchers who have already developed support approaches, as they have revealed that not all students have benefited equally from the given support (Kim *et al.*, 2022; Petritis *et al.*, 2022; Stowe and Cooper, 2019). Accordingly, in order to appropriately address individual challenges, these challenges must first be diagnosed. In this regard, Kuhn (1992) stated that challenges should be identified in real-life contexts, especially in work contexts, because this is where people are usually confronted with the greatest challenges, which also makes more complex thought processes prevailing (Kuhn, 1992). Therefore, it has already been noted that the work context of a chemist is characterized by the formation of hypotheses and the weighing of alternatives in the form of arguments (Kuhn, 1992; Lombardi *et al.*, 2016; Longino, 1990; Osborne and Patterson, 2011). Accordingly, the diagnosis of students' challenges in building arguments and the subsequent adaptive support of these challenges should take place through the weighing of alternatives because "*only by considering alternatives […] one can begin to achieve any certainty about what is*" (Kuhn, 1992, p. 164). The aforementioned theoretical considerations form the basis of our work to diagnose students in building arguments on alternative reaction pathways in organic chemistry and to adaptively support them in arguing and using concept knowledge.

1.2 Research Objectives

The linkage of building reasonable arguments and applying appropriate scientific principles plays a central role in enabling students to participate in scientific discourse (Bricker and Bell, 2008; Driver *et al.*, 2000). Therefore, in order to foster students in participating, the need occurs to investigate how students build arguments, *i.e.*, to diagnose students' approaches in general as well as the challenges students experience, and to support them individually. Accordingly, the objectives of this dissertation were to investigate how students build arguments when judging the plausibility of alternative reaction pathways in organic chemistry. Based on these findings, it was the goal to design an adaptive scaffold to provide students with individual support in building arguments and applying concept knowledge. Therefore, this dissertation is guided by the following two hypotheses:

1) A qualitative analysis based on theoretical considerations in argumentation research can provide new insights into how students build arguments on alternative reaction pathways in organic chemistry and the challenges they experience.
2) Adaptive scaffolding based on the diagnosed challenges can support students regarding their individual needs in their argumentation skills and use of concept knowledge.

1.3 Methods

In this dissertation, two studies were conducted to investigate (1) how students build arguments when judging the plausibility of alternative reaction pathways in organic chemistry and (2) whether an adaptive scaffold, designed on the basis of the results of the first study, can support students in building arguments and using concept knowledge while justifying the plausibility of alternative reaction pathways. Figure 2 summarizes the complete research design, which is described in the following.

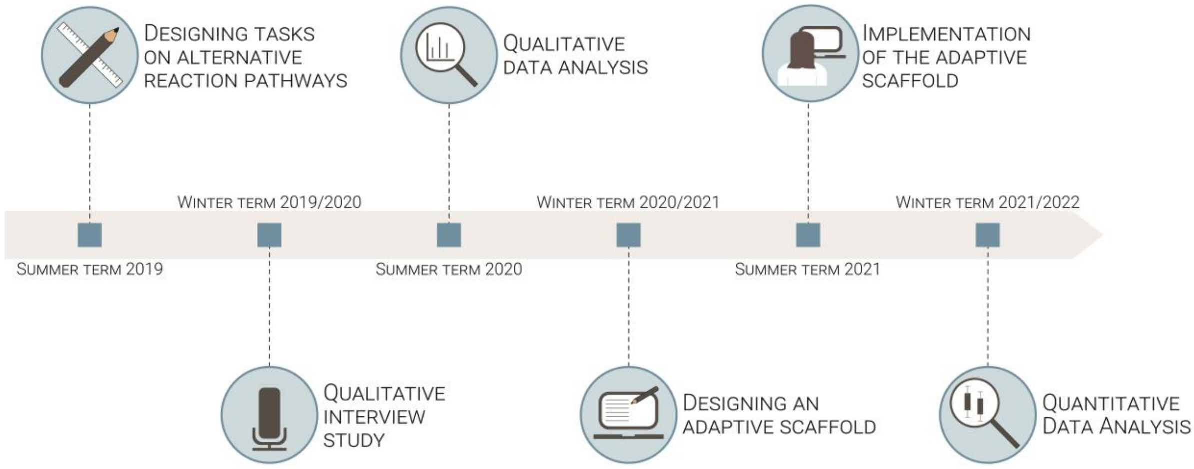


Fig. 2: Research design of the complete research project.

At the beginning of the research project, a task design consisting of alternative reaction pathways took place. Therefore, two tasks were designed intending to cause a cognitive dissonance by judging the plausibility of alternative reaction pathways in organic chemistry. Both tasks consisted of four subtasks, while the first task was centered on the reaction of 4-chlorobutanol with hydroxide, the second task was centered on the reaction of methyl acetate with diisopropylamide. In the first subtask, students were asked to build a product for a typical organic chemistry reaction. In the second subtask, students were again asked to build a product for a reaction that differed from the previous reaction only by one additional surface feature. Subsequently, in subtask 3, the students were given five alternative reaction products on product cards for the reaction from subtask 2, for which they were to judge the plausibility. It is to notice that not all of the provided alternative reaction products were correct. Finally, the students had the opportunity to defend or revise their product formed in subtask 2. A more detailed description of the tasks can be found in chapter 2. The original tasks and their sample solutions are provided in the Appendix (6.2.1 and 6.2.2).

Study 1 was a qualitative interview study conducted at the Justus-Liebig-University Giessen in October and November 2019. The think-aloud interviews were conducted using a semi-structured interview guide provided in the Appendix (6.2.3) (Bernard and Bernard, 2013). Twenty-nine students from the course "Organic Chemistry 3 - Catalysis and Synthesis" were recruited for the interviews on a voluntary basis. All students were enrolled in the previous organic chemistry courses (OC I and OC II), which meant that the content of these courses was assumed to be students' prior knowledge. Content of the previous courses included reactivity of functional groups, typical reaction mechanisms, and structure-property relationships, among other topics. At the beginning of the individual interviews the students indicated that they agreed to allow their data to be analyzed and used for research purposes. They completed a demographic questionnaire and were informed about their rights. The demographic questionnaire and consent form are provided in the Appendix (6.1.1 and 6.1.2). During the interviews, students completed the two tasks, which were previously mentioned. The collected data includes the demographic questionnaire, the completed worksheets, and video and audio recordings. Since only native German speakers participated in the study, interviews were conducted in German, transcribed verbatim, and analyzed. Corresponding quotations were translated into English and are listed in the Appendix (6.4). The data analysis took place using the software MAXQDA and consisted of three sequential steps, which are described in more detail in chapter 3. In the first step, the argumentation patterns were identified for each student and for each product card (see Appendix 6.2.4 for students' argumentation patterns). This involved using the CER model to determine claims, pieces of evidence, and reasoning statements uttered by the students (McNeill and Krajcik, 2012; McNeill *et al.*, 2006; Osborne and Patterson, 2011). In the second step, data-based reasoning categories were built from students' reasoning statements. In the third step, the number of pieces of evidence and reasoning statements and the ratio of reasoning to evidence were used to determine three argumentation approaches into which students were assigned.

Study 2 was a quantitative study that took place online at Tufts University in Boston (USA) in April and May 2021 and received Institutional Review Board approval for human subjects research (STUDY00001480). To conduct this study, a two-part adaptive scaffold was developed using the application Qualtrics, which consisted of a diagnostic scaffold that all students worked on equally and four adapted scaffolds, on which students worked on one of the four adapted scaffolds depending on their performance in the diagnostic scaffold two weeks later. Sixty-four students from the Organic Chemistry II course participated in the study on a voluntary basis. However, students who did participate received extra credit for their participation, which made up 1% of their final course grade. Again, the content of the previous organic chemistry course was assumed to be students' prior knowledge. Before students could work on the diagnostic scaffold, they completed a consent form and a demographic questionnaire, which are provided in the Appendix (6.1.3). The scaffolds can also be found in the Appendix (6.3) and are presented in more detail in chapter 4 and chapter 5. The collected data includes the demographic questionnaire and students' answers of the diagnostic scaffold and the adapted scaffolds. The data analysis consisted of three parts. The first part of the data analysis took place immediately after the diagnostic scaffold. Here, students' answers were analyzed to determine an argumentation score and a concept knowledge score for each student. In the second step, students were assigned to one of four adapted scaffold groups based on their argumentation score and concept knowledge score (see chapter 4). After students completed the adapted scaffold, the third and final step of the data analysis took place. Here, students' answers were scored again, giving each student a second argumentation score and concept knowledge score. Based on the scores, quantitative tests were used to determine whether the adaptive scaffold helped students to build arguments and to use concept knowledge. For this purpose, the software R 4.0.4 (R Core Team, 2021) was used with the packages ggplot2 (Wickham, 2016), psych (Revelle, 2021), rcompanion (Mangiafico, 2021), stats (R Core Team, 2021), base (R Core Team, 2021), and graphics (R Core Team, 2021). A detailed description of the methods can be found in chapter 4.

1.4 Results and Discussion

In the following chapter, the two preceding hypotheses will be answered. The first part deals with the diagnosis of students when building arguments on alternative reaction pathways. The focus is placed on a classification of argumentation approaches, which allows the analysis of students' individual challenges and their subjective feedback. The second part deals with the support of students based on their individual needs through the adaptive scaffold. Thereby, the effectiveness of the scaffold along with the personal evaluation of the students is discussed. In both hypotheses, the main findings of the two studies are summarized and their contribution to the literature is provided. Student quotes are from the German interview study and the U.S. investigation and were used in the publications unless noted otherwise (see chapter 2-5).

1.4.1 New insights into how students build arguments on alternative reaction pathways in organic chemistry and the challenges they experience

To build well-grounded arguments in chemistry, it is important to be able to structure arguments as well as to activate and apply prior knowledge (Choi *et al.*, 2013; Sandoval and Millwood, 2005; von Aufschnaiter *et al.*, 2008). Teaching concept knowledge is a central component of organic chemistry (Cooper, 2015; Stowe *et al.*, 2021). In contrast, building arguments and weighing alternative reaction pathways are competencies that are new to most students. For this reason, the analysis of argumentation on alternative reaction pathways focused initially on the structuring of arguments and the general use of chemical concepts and not on the correct use of scientific principles. In its simplest form, the structure of an argument consists of a claim that is supported with evidence and justified with reasoning (CER model) (McNeill and Krajcik, 2012; McNeill *et al.*, 2006). Thus, the *claim-evidence-reasoning* model offers a promising approach to characterize how students build arguments, which has already been successfully applied by several researchers (Bodé *et al.*, 2019; Deng and Flynn, 2021; Luo *et al.*, 2020; Petritis *et al.*, 2021; Walker *et al.*, 2019).

To gain new insights into the process of building arguments on alternative reaction pathways and to be able to support students in a targeted manner, the use of the three argument components, *i.e.*, claim, evidence, and reasoning, was analyzed. Three argumentation approaches were identified, which build a foundation to illustrate the results, namely the reasoning-based argumentation approach, the based-on-both argumentation approach, and the evidence-based argumentation approach.

How do students use evidence and reasoning in their argumentation on alternative reaction pathways?

To investigate how students use evidence and reasoning in their argumentation, "argumentation trees" (see Figure 4) were created from students' argumentation patterns based on the interviews of each student and each of the five alternative reaction pathways. Complete argumentation patterns of all students are provided in the Appendix (6.2.4). By analyzing the frequency of pieces of evidence and reasoning statements and the ratio of reasoning to evidence, three argumentation approaches were characterized (see Figure 3). Students in the reasoning-based argumentation approach (RB approach) justified each piece of evidence on average with at least one reasoning statement (ratio > 1). Students in the based-on-both argumentation approach (BoB approach) supported their claims also with evidence and reasoning, but less frequently than students in the RB approach (ratio ~ 1). Students in the evidence-based argumentation approach (EB approach) provided the least evidence and reasoning statements to support the claim, however, not all pieces of evidence were justified with reasoning (ratio ≤ 1).

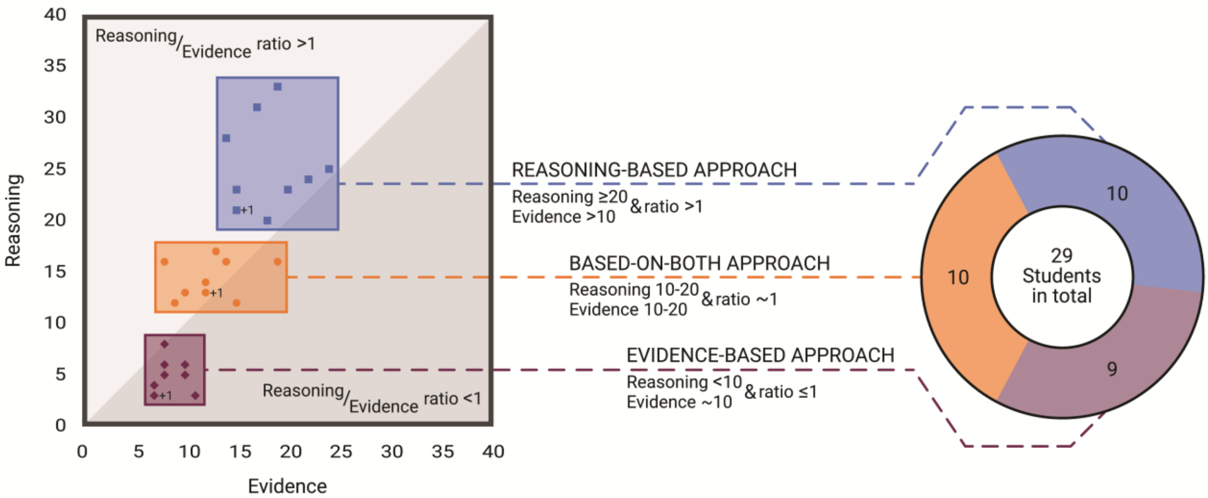


Fig. 3: Classification of the three argumentation approaches related to the frequency of evidence and reasoning and the ratio of reasoning to evidence.

In addition to the frequency and ratio of the claims' justification, the study also examined the ways in which students used reasoning. The analysis revealed eight data-based reasoning categories: *Electronics*, *Energetics*, *Kinetics*, *Spatial Arrangement*, *Strength*, *Stability*, *Analogies*, and *Conditions* (a detailed description of the reasoning categories can be found in chapter 3). Students in the RB approach provided the most statements in each reasoning category, whereas students in the EB approach provided the least. However, the percentage distribution of reasoning categories was similar for all three approaches (see Table 1).

Table 1: Frequency (number of appearances, left side of the column) and percentage (right side of the column) of reasoning statements in the eight data-based reasoning categories for the reasoning-based argumentation approach (RB), for the based-on-both argumentation approach (BoB), and for the evidence-based argumentation approach (EB).

	Electronics		*Energetics*		*Strength*		*Spatial Arrangement*		*Stability*		*Analogies*		*Kinetics*		*Conditions*	
RB	98	39%	42	17%	36	14%	30	12%	26	10%	8	3%	7	3%	5	2%
BoB	61	43%	24	17%	14	10%	14	10%	19	13%	5	4%	3	2%	2	1%
EB	15	40%	6	16%	2	5%	8	22%	4	11%	1	3%	1	3%	0	0%

The following three examples (see Figure 4) serve to illustrate the three reasoning approaches and the implementation of the reasoning categories. All students claimed that THF is a plausible product of the reaction of 4-chlorobutanol with hydroxide. The quotes are not part of the publications.

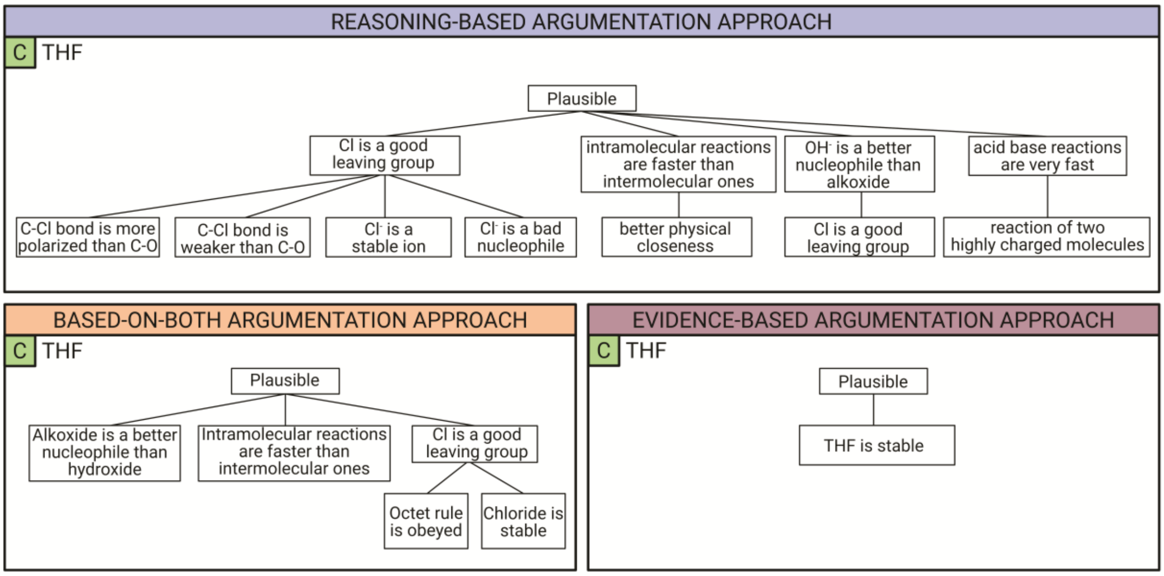


Fig. 4: Example argumentation patterns for the alternative reaction pathway to THF: (top) argumentation pattern of Charlie (reasoning-based argumentation approach); (bottom left) argumentation pattern of Sydney (based-on-both argumentation approach); and (bottom right) argumentation pattern of Amber (evidence-based argumentation approach).

Charlie, a student from the reasoning-based argumentation approach, used four pieces of evidence and seven reasoning statements to justify his claim. Each piece of evidence was justified with at least one reasoning statement. Charlie provided reasoning statements of five reasoning categories, such as *Electronics* ("*C-Cl bond is more polarized than C-O*" and "*reaction of two highly charged species*"), *Energetics* ("*C-Cl bond is weaker than C-O*"), and *Strength* ("*Cl^- is a bad nucleophile*"). Sydney, a student from the based-on-both argumentation approach, built three pieces of evidence for the claim, one of which was justified with two reasoning statements that belonged to the reasoning categories *Stability* ("*Cl^- is stable*") and *Electronics* ("*octet rule is obeyed*"). Amber, who was assigned to the evidence-based argumentation approach, supported her claim with one piece of evidence but no further reasoning statements, which is why no reasoning category was applied. A detailed description of the argumentation approaches and further examples can be found in chapter 3.

Even though all three students came to the same conclusion at the end, namely that THF is a plausible reaction product, their argumentation differed significantly. Based on the examples, it is evident that the number of evidence and reasoning statements can increase the depth and breadth of an argument, and thus its quality, as the claim is justified with multiple chemical concepts. This is consistent with the literature, which states that the quality of an argument is determined by a close linkage between argument components, *i.e.*, evidence and reasoning, and the required concept knowledge (Choi *et al.*, 2013; Sandoval and Millwood, 2005; von Aufschnaiter *et al.*, 2008), as evidence and reasoning give weight to the argument. However, quantity does not necessarily equal quality regarding the number of argument components (Cross *et al.*, 2008). Moreover, the use of different reasoning categories in all three reasoning approaches is also a promising finding, as prior studies have found that students experience challenges integrating multiple variables into their decision-making process (Kraft *et al.*, 2010; Sevian and Talanquer, 2014; Talanquer, 2006; Weinrich and Talanquer, 2015). However, considering multiple variables when building arguments and weighing them in the decision-making process supports students not only in improving the quality of their arguments, but also in better understanding the reactions that occur as they elaborate on the causes of structural changes (Caspari *et al.*, 2018; Cooper *et al.*, 2016; Grove *et al.*, 2012; Watts *et al.*, 2020). However, it has also been noticed, particularly among students in the evidence-based argumentation approach, that they experienced challenges when using evidence and especially reasoning, as illustrated by the example of Sonia. She had difficulty predicting the mechanism for the formation of THF, but had already claimed THF as the most plausible product of the reaction.

Interviewer*: "We can do it differently. When you see the product, would you describe the product as plausible in principle?"*

Sonia*: "Yes, but I don't know why. My feeling tells me that again, I don't know. It looks so right somehow."*

Interviewer*: "And are there any factors that help you determine why you think the product could be right?"*

Sonia*: "We once had a similar task in an exercise and I don't know, when I saw that, it kind of clicked in my head, somewhere in the back corner where I thought, I think that's it."*

Sonia could not justify her claim with evidence nor reasoning, but relied solely on her gut feeling. Even though her gut feeling did not let her down in this particular case, it might not be satisfactory for a chemist to be unable to provide any justification for a decision. Relying on a personal experience such as gut feeling rather than providing a well-grounded argument is a well-known phenomenon among learners (Hogan and Maglienti, 2001; Sadler and Zeidler, 2004). This might relate to how learners handle uncertainty with concept knowledge. However, it is also possible that a certain phenomenon has never been questioned by the learner, which made it seem unnecessary to build a chemically sound argument, as Amber shows when judging the plausibility of a carbanion. The quote was not used in the publications.

Interviewer: *"Do you think the product (carbanion) is plausible?"*

Amber: *"Actually not because here is a negative charge at the Cl."*

Interviewer: *"And why does that bother you?"*

Amber: *"I think I have never seen that before."*

Amber saw no need to build a justification for her claim, she relied on her experience instead. This approach is also well known in the literature. McDonald (2010), in her study of the influence of nature of science and argumentation instruction on views of nature of science, interviewed a student named David who responded to a question about atomic structure as follows.

David: *"Well, that's one of those silly things that I can't explain … Right now I've never seen one, I put my faith in science … I suppose it's part of my personality, my psyche, I'm just scientifically inclined. And I blindly believe the scientists." (McDonald, 2010, p. 1161)*

The reason why students like Sonia, Amber, and David relied on their gut feeling or experience instead of building well-grounded arguments could be that they do not realize that claims are not absolute facts but need to be questioned and thus need to be further justified by evidence and reasoning (McDonald, 2010). This leads to them being less engaged in the argumentation process (Kuhn and Reiser, 2006).

Do students change their initial claims after building arguments on alternative reaction pathways?

Students in study 1 and study 2 received the same two reactions and built arguments on the same alternative reaction pathways, although in study 2 they were each reduced by one alternative reaction pathway (carbanion and N-O attack). At the beginning of the interview/scaffold, the students of both studies made an initial claim about the product of the given reaction. After students had built arguments on the alternative reaction pathways, they had the opportunity to defend or revise their initial claim. It is to mention that students in study 1 built their arguments verbally, while students in study 2 built them in writing. Figure 5 summarizes the results of both studies with respect to claim change. Detailed descriptions can be found in chapters 2, 3, and 5.

In the first reaction of 4-chlorobutanol with hydroxide, students in both studies predicted mainly the incorrect reaction product 'Diol' when making the initial claim (study 1 58.5%, study 2 81%, see Figure 5). After building arguments on alternative reaction pathways, the number of the correctly predicted reaction products 'THF' and 'Alkoxide' increased tenfold for students in study 1 (from 5% to 51%) and doubled for students in study 2 (from 16% to 30%). A similar trend is seen for the second reaction of methyl acetate with diisopropylamide. Students in study 1 achieved an increase in the correct products of 'Methyl acetoacetate' and 'Enolate' from 50% to 83% and students in study 2 from 23% to 63%. Here, it becomes apparent that students in both studies increased in their number of correctly predicted reaction products.

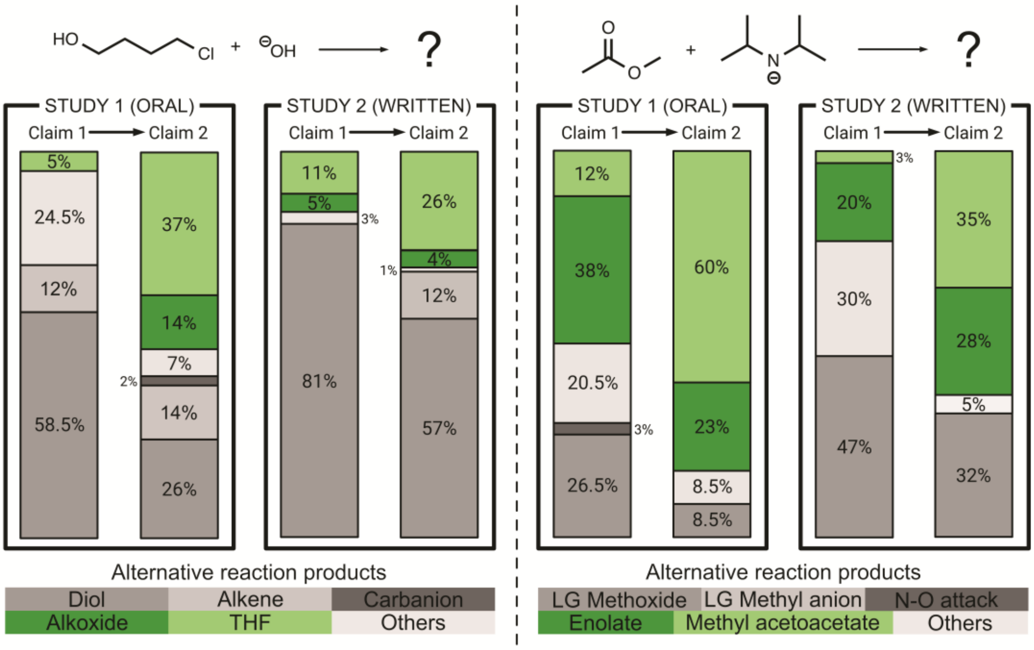


Fig. 5: Total of all "most plausible" claims made by the students of both studies. Students in study 1 built arguments orally, whereas students in study 2 built written arguments. Students were able to claim more than product as plausible. "Others" refers to incorrect products such as intermediates.

As an example, Charlie's moment of claim change is presented. He is a student from study 1 (reasoning-based argumentation approach). Charlie originally claimed the 'Diol' as the reaction product. After the argumentation process he decided to revise his initial claim by predicting 'THF' with 'Alkoxide' as the precursor as most plausible product of the reaction.

Charlie: *"I don't think it's not plausible (he refers to THF). I wouldn't have intuitively guessed that something like this (THF) would happen if I am honest. But I would have taken an easier way and said it is a normal nucleophilic substitution, at this point, if I look at it that way I would not necessarily say it (THF) could be a by-product of what is being created. But the main product, the more I think about it, the more fascinating I think it is."*

Interviewer: *"Why?"*

Charlie: *"Because I really didn't think that this (formation of THF) was a possibility. Well, I didn't think about it at all, I even think that what was there before (nucleophilic substitution) is wrong, because I just saw that I have a nucleophile and a good leaving group and nucleophilic substitution, but I didn't see that you can build another nucleophile (alkoxide)."*

However, several students also stayed with their incorrect initial claims. The most common reason for sticking with their incorrect initial claim in both studies might be that students could not activate the necessary prior knowledge and thus could not assess, for example, whether an acid-base reaction was occurring because they did not know the corresponding pK_a values. In study 1, this occurred particularly frequent in the evidence-based argumentation approach. However, it can also be seen that students in study 1, who built arguments verbally, were more likely to make a claim change than students in study 2, who built written arguments. That students experience challenges changing their claim when building written arguments is consistent with the literature (Chang and Chiu, 2008; Luo *et al.*, 2020; Novak and Treagust, 2018; Walker *et al.*, 2019). Here, it has been noticed that it is difficult to change a person’s point of view (Novak and Treagust, 2018), especially in a short period of time (Chang and Chiu, 2008). Instead, students often stick to their claim even when facing critique and do not consider alternatives in their decision-making process (Walker *et al.*, 2019). Walker *et al*. (2019) also noticed this in the follow-up discussions of their study, in which students continued to stick to their claim even when they received strong counterarguments. Possible reasons for students sticking with their incorrect claims, in addition to a lack of concept knowledge (Novak and Treagust, 2018), may be that students perceive a claim change as a sign of failure of their previous work (Walker *et al.*, 2019) or that new data are ignored that could change one’s claim (Novak and Treagust, 2018). This was also noticeable in study 2. Mary, when arguing on the alternative reaction pathway to ‘THF’, stated that she thought ‘THF’ is a plausible reaction product and built the following argument. The argument is not a highlighted example in the publications.

Claim: THF is a plausible reaction product

Evidence 1: The 4-chloro-1-butanol hydroxyl group can be reasonably deprotonated by hydroxide anion.

Reasoning 1: The acidic proton of the reactant's OH group could be deprotonated by hydroxide, as hydroxide is a fairly strong base. The conjugate acid (water) is very stable, and the conjugate base (an alkoxy anion) is not wildly unstable.

Evidence 2: The pictured product could be reached by an intramolecular S_N2 reaction between the deprotonated OH and the electrophilic carbon involved in the C-Cl bond.

Reasoning 2: The C of the C-Cl bond is electrophilic due to the inductive effect of the adjacent electronegative chlorine, so it is vulnerable to nucleophilic attack. The alkoxy anion at the other end of the molecule (following deprotonation) could then attack this electrophilic carbon.

Evidence 3: Chloride is a good leaving group.

Reasoning 3: Chloride is a weak base and electronegative, making it well able to stabilize a negative charge.

Mary provided a detailed justification for her claim without finding any counterarguments for the formation of ‘THF’. Nevertheless, regarding the question of whether she would like to stay with her initial claim or revise it, she decided to stick to her initial claim. In line with the possible reason of Novak and Treagust (2018), Mary seemed to be able to build a new justification, but this new data appeared not sufficient enough to her to change her claim.

How do students reflect their problem-solving approach when building arguments on alternative reaction pathways?

It was part of the interviews in study 1 for students to reflect on their problem-solving approaches when building arguments on alternative reaction pathways and to provide feedback. Thereby, students provided promising insights. Figure 6 illustrates an example from each of the three argumentation approaches. The quotes from Lily and Amber are not part of the publications. Further examples of feedback from students in study 1 can be found in chapter 2.

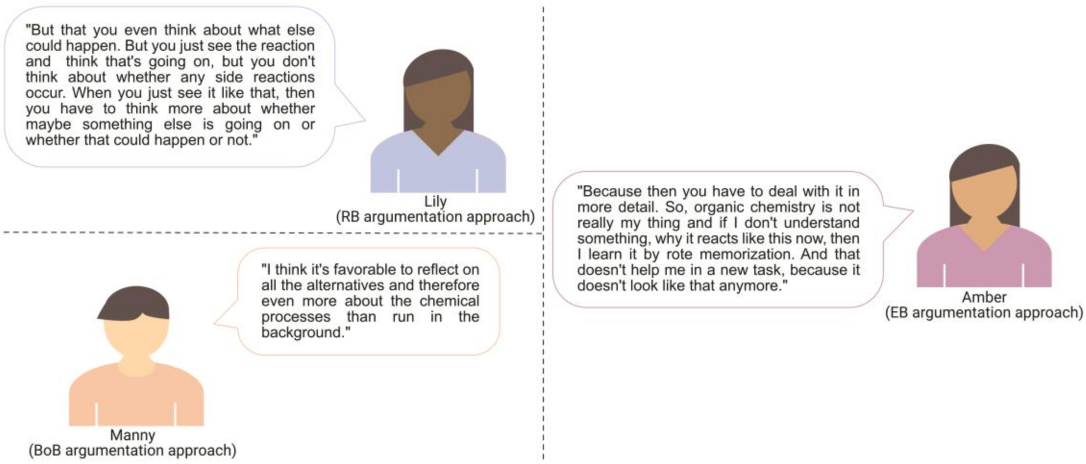


Fig.6: Examples of students' feedback in the interviews of study 1. Each argumentation approach is represented with one example: Lily from the reasoning-based argumentation approach, Manny from the based-on-both argumentation approach, and Amber from the evidence-based argumentation approach.

In general, students were able to reflect on their problem-solving approaches and assess the productivity of their approaches. In addition, the task format received positive feedback, but also resulted in certain students feeling uncertain, possibly due to the fact that students were not used to the task format or they experienced cognitive overload (Sweller, 2003).

When reflecting on the problem-solving approach, Lily and the interviewer talked about whether building arguments on implausible reaction pathways also supports the use of chemical concepts. Lily affirmed this and argued that she does not usually think about side reactions or byproducts. However, when building arguments on implausible reaction pathways, she thinks more about what else might happen. This is consistent with the findings of Popova and Bretz (2018), who found in their research on reaction coordinate diagrams that byproducts are not considered important, but that students focus mainly on the main product of a reaction. One reason for this could be that students perceive reaction equations and mechanisms literally (Popova and Bretz, 2018) and so they are also sometimes unaware that a reaction mixture does not consist of only one molecule of reactant. Students are thus not necessarily aware that reaction equations or mechanism do not reflect reality (Bodner and Domin, 2000), which can be attributed to the fact that students tend to build static concepts compared to experts (Bhattacharyya and Bodner, 2005; Bodner and Domin, 2000). Building arguments, especially about implausible reaction pathways, made Lily question her problem-solving approach and broaden her horizon towards reaction pathways and associated side reactions. In Manny's reflection on the benefits of considering alternative reaction pathways after building arguments, he expressed that it is beneficial to reflect on alternatives because it also makes one think more about the chemical processes that are going on in the background. Manny and Lily both expressed that they found reflecting on alternative reaction pathways helpful in being able to activate their concept knowledge as well. Therefore, it can be surmised that they focused less on a product and instead considered the process by weighing chemical concepts. This approach is similar to that of an expert as problems are explored (Flynn, 2011), prior knowledge is organized (Kraft *et al.*, 2010), and information is analyzed (Cartrette and Bodner, 2009). Furthermore, in considering the process, the focus is placed more on weighing underlying chemical concepts and considering and reasoning about mechanisms rather than remembering facts and 'playing with puzzles' (Bhattacharyya and Bodner, 2005; Caspari *et al.*, 2018; Grove *et al.*, 2012; Talanquer, 2013; Watts *et al.*, 2021). Remembering facts, however, was always Amber's problem-solving approach, as she expressed when reflecting at the end of the interview. By her own admission, organic chemistry was never 'her thing', which is why she focused on memorization. However, she was also able to critically reflect on her problem-solving approach, noting that memorization did not help her in solving new problems. That students often use memorization rather than understanding underlying concepts is a well-known approach in organic chemistry. Similar to Amber, students try to recognize patterns, which leads them to focus more on explicit features, neglecting the implicit properties of molecules (Anzovino and Bretz, 2015; Cooper *et al.*, 2013; Gigerenzer and Gaissmaier, 2011; Graulich, 2014; Morewedge and Kahneman, 2010; Talanquer, 2017). It is therefore encouraging that after building arguments on alternative reaction pathways, students were critical of their problem-solving approaches and experienced that memorization related to these task formats is unproductive.

1.4.2 New insights into how adaptive scaffolding can support students regarding their individual needs in structuring their arguments and use of concept knowledge

Subsection 1.4.1 highlighted how students build arguments on alternative reaction pathways and the challenges they experience in doing so. In the subjective evaluation of the students, it was found that they often relied on their memorized knowledge. However, by weighing alternative reaction pathways, many students were also able to activate and apply their concept knowledge. In the analysis, the students' assessment was confirmed. Here, there were students, especially from the reasoning-based argumentation approach, who supported their claims with many pieces of evidence and reasoning statements, but also students, especially from the evidence-based argumentation approach, who could hardly build a justification for their claims. Students experienced the greatest challenges in building reasoning statements, which is congruent with the literature findings (Deng and Flynn, 2021; Petritis *et al.*, 2021; Stanford *et al.*, 2016). Analysis of the reasoning categories also revealed that reasoning statements were mainly built from the category *Electronics*, although multivariate use of concepts would be desirable (Kraft *et al.*, 2010; Sevian and Talanquer, 2014; Talanquer, 2006; Weinrich and Talanquer, 2015), which several students have already applied. During the interviews, it was found that many students had difficulties to activate the necessary prior knowledge to be able to build an appropriate justification. This observation was confirmed in the analysis of changing initial claims, where several students referred to personal views instead of supporting their justification with scientific principles, which is consistent with results from the literature (Hogan and Maglienti, 2001; Kuhn and Reiser, 2006; McDonald, 2010).

To address students' individual needs, in study 2 an adaptive scaffold was designed based on four instructional approaches (Kang *et al.*, 2014; Walqui, 2006). A more detailed description of the scaffold design can be found in chapter 5. Moreover, the scaffolds are provided in the Appendix (6.3.1-6.3.5). Thus, the following subsection focuses on students' individual challenges in argumentation on alternative reaction pathways regarding the scaffold and the effectiveness of the adaptive scaffold.

Which challenges do students experience in their argumentation on alternative reaction products?

At the beginning it was noted that scaffolding provides a promising approach to support students with their experienced challenges. Most common are so-called *hard* or *fixed* scaffolds, which provide all students with the same level of support (Azevedo *et al.*, 2004; Chen, 2014). Thereby, the question arises whether a scaffold is equally suitable for all students, which has already been answered negatively in several studies. Reasons for this include that individual student challenges are not addressed (Azevedo *et al.*, 2004; Chen, 2014), students are usually not diagnosed (Chen, 2014; Greene and Land, 2000; Puntambekar and Hubscher, 2005), and that scaffolds often only have benefit for single groups, such as low-performers (Kranz *et al.*, 2022). The diverging success through fixed scaffolds among students relates to the fact that students should work in the zone of proximal development (Vygotsky, 1980). This means that each student has a zone in which he or she cannot complete tasks independently but only with additional guidance (Kang *et al.*, 2014; Lajoie, 2005; Pea, 2004; Wood *et al.*, 1976), thus challenging but not overwhelming students. However, as students experience individual challenges and differ in their prior knowledge, scaffolds should be tailored to students' needs (Pea, 2004; Van de Pol *et al.*, 2010; Vygotsky, 1980). Adaptive scaffolds offer an opportunity in this regard, as continuous diagnosis identifies students' challenges (Kang *et al.*, 2014; Lin *et al.*, 2012) and thus adapts the scaffold to students' challenges (Azevedo *et al.*, 2004; Azevedo *et al.*, 2008; Chen, 2014; Kang *et al.*, 2014; Van de Pol *et al.*, 2010).

Therefore, to support students based on their individual strengths and challenges, these must first be diagnosed. For this reason, the adaptive scaffold consisted of a diagnostic scaffold as a pre-measure in which students already received support on the structure of an argument and applying concept knowledge while being diagnosed. Based on the literature findings and the results of study 1, it was hypothesized that students would experience challenges in (1) structuring arguments, (2) applying concept knowledge, (3) both areas, or (4) building multivariate arguments. Two scoring systems (argumentation and concept knowledge) were designed to diagnose students' individual challenges. The scoring system can be found in chapter 4 and with the following DOI 10.17605/OSF.IO/4ZPN9. Students were assigned to one of four adapted scaffolding groups based on the scoring results of the diagnostic scaffold (pre-measure), *i.e.*, the challenges presented earlier, providing support in the area in which they experienced the greatest challenges. After passing the adapted scaffolds, students' answers were scored again (post-measure). Figure 7 illustrates four student examples from the diagnostic scaffold, each of which was assigned to one of the four adapted scaffolding groups.

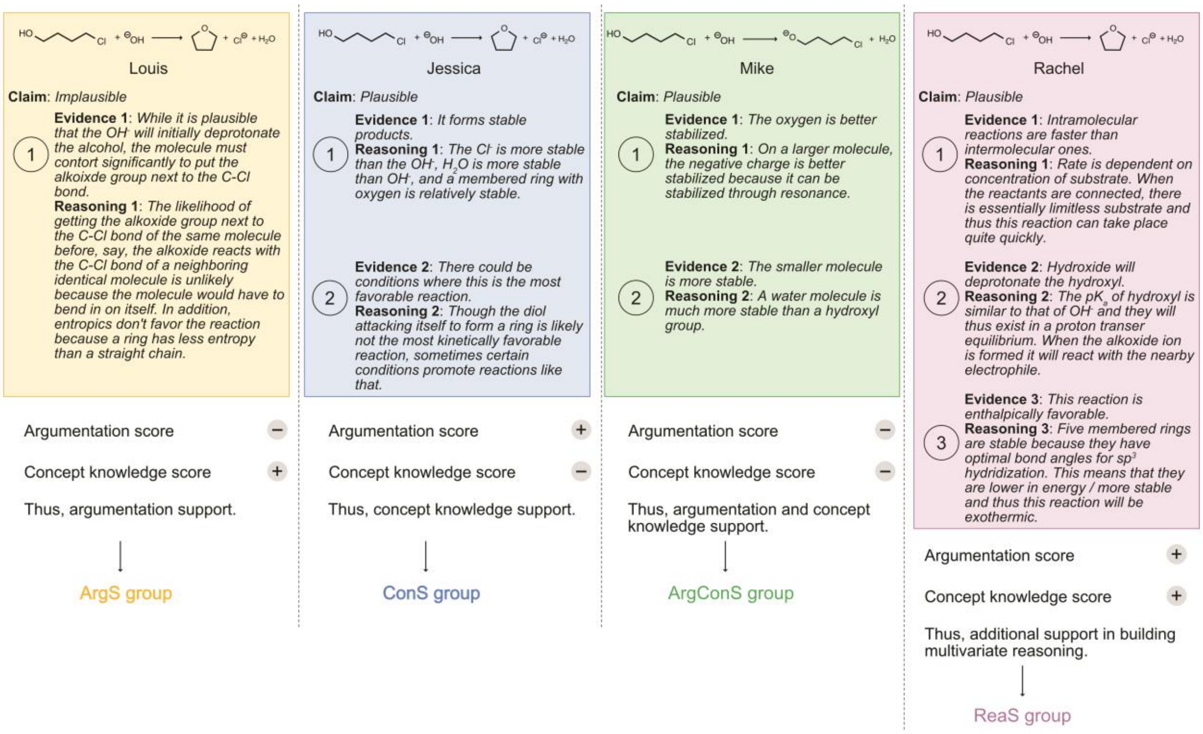


Fig.7: Exemplary arguments of four students who were assigned to the four different adapted scaffolding groups after the diagnostic scaffold. Students' assignments were based on their argumentation score and concept knowledge score.

Louis experienced challenges in separating the argument components, *i.e.*, evidence and reasoning, when building arguments on alternative reaction pathways. For example, in Reasoning 1 (see Figure 7, left column), he simply repeated the evidence on the one hand and mentioned another evidence with an associated reasoning statement regarding entropy on the other hand. For this reason, Louis received support in the adapted scaffold for the structure of argument components (ArgS group). Jessica, on the other hand, was able to separate argument components, but had difficulty supporting evidence and reasoning with scientific principles. Therefore, she received support in applying concept knowledge (ConS group). Mike experienced challenges in both, the structure of argument components and applying concept knowledge and thus received support in both areas (ArgConS group). This is evident in Figure 7, as Mike, for example, reasoned incorrectly in Reasoning 1, did not make it apparent in Evidence 2 which molecule is more stable than the other, and Reasoning 2 is not a justification of Evidence 2, as the evidence is only concretized. Rachel is already conspicuous by the multiplicity and length of her argument components in Figure 7. Moreover, she justified her claim with multiple chemical concepts that are justified precisely. To further support Rachel as well, the focus of the adapted scaffold was on justifying the claims with multivariate reasoning (ReaS group).

To what extent does the adaptive scaffold close the gap in students' performance?

In the quantitative analysis of the scoring results, Kruskal-Wallis tests were performed for the argumentation score and the concept knowledge score, respectively, to determine group differences according to the diagnostic scaffold (pre-measure) and according to the adapted scaffolds (post-measure). In case of significant results, post-hoc comparisons with Wilcoxon rank-sum tests and Bonferroni adjusted *p*-values were performed (Field *et al.*, 2012; Rosenthal, 1991). The correlation coefficient *r* was defined as .10 ≤ *r* ≤ .30 as small effect, .30 ≤ *r* ≤ .50 as medium, and *r* ≥ .50 as large (Cohen, 1992). A more detailed description of the methodology and statistical results can be found in chapter 4. Figure 8 presents a summary of the group comparisons.

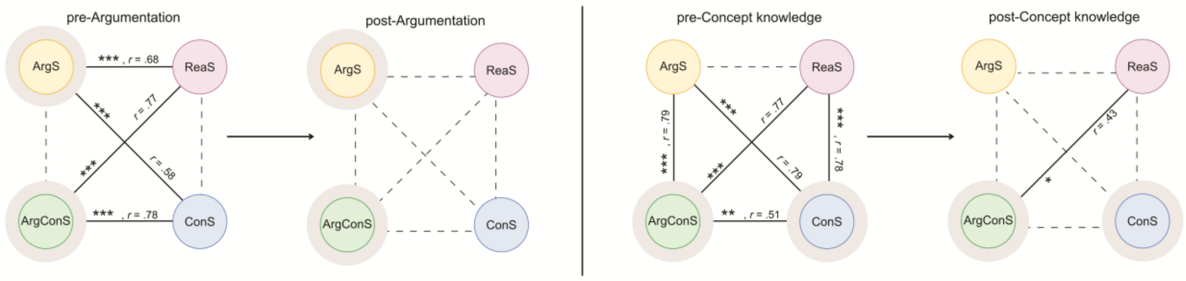


Fig.8: Differences between the group scaffolding groups before (pre) and after (post) the adapted scaffolds. In case of non-significant differences, the lines are dashed. Groups who received an adapted scaffold in the respective area of support are highlighted with a gray background circle. Significance levels of the group comparisons are indicated (* *p* < .05; ** *p* < .01; *** *p* < .001).

It should be noted in advance that the students were not yet assigned to an adapted scaffolding group in the pre-measure. However, since they subsequently received support from one of the four adapted scaffolds on the basis of the scoring results in the pre-measure, the assignment of the students is already presented in the pre-measure for a better illustration of the pre- and post-measure comparison. Specifically, this means that Mike (see Figure 7), for example, received a low argumentation score and a high concept knowledge score in the pre-measure and was therefore assigned to the ArgS group. On Figure 8, he is thus also part of the ArgS group in the pre-measure.

The left side of Figure 8 shows the comparison of the argumentation scores. In the pre-measure, four of six groups differed significantly from each other with a large effect (*r* = [.58;.78]). The two groups that received additional argumentation support after the pre-measure (ArgS and ArgConS group) did not differ significantly from each other in their argumentation score. The two groups that did not receive support in argumentation after the pre-measure (ConS and ReaS group) also did not show significant differences in the argumentation score. This means that the division of the groups based on the argumentation score can be regarded as successful. In the post-measure, there were no significant group differences, so the gap between the groups regarding the argumentation score could be closed. In terms of concept knowledge, five of six groups differed significantly from each other with large effects (*r* = [.51;.79]) in the pre-measure. Only the two groups that did not receive additional support in applying concept knowledge after the pre-measure (ArgS and ReaS group) showed no significant difference. Therefore, one may pose the question why the ArgConS group and the ConS group both received support in concept knowledge after the pre-measure, even though they were noteworthy different from each other in the pre-measure. First, the concept knowledge scores of both the ArgConS group and ConS group were significantly lower in contrast to the two groups that did not receive additional support in concept knowledge (ArgS and ReaS group). Moreover, in the qualitative analysis of the arguments built, it is noticeable that the applied concept knowledge of the ArgConS group was noteworthy lower than that of the ConS group. For this reason, it is legitimate that both groups received additional support in applying concept knowledge. In the post-measure, only the ReaS group and the ArgConS group still differed significantly from each other, which is also due to the fact that the ReaS group obtained the highest scoring results in concept knowledge and the ArgConS group obtained the lowest. Nevertheless, the gap in concept knowledge could be narrowed between all groups.

When studying the four adapted scaffolding groups, the question may arise to what extent argumentation and concept knowledge can be fostered separately from each other. At the beginning, it was already mentioned that argumentation and concept knowledge are closely linked. However, whether there is a relationship between the learning success of argumentation and concept knowledge is a matter of disagreement among researchers, as there are voices in favor (Acar, 2008; Bell and Linn, 2000; Noroozi *et al.*, 2017; Songer and Gotwals, 2012) and there are voices opposed (de Lima Tavares *et al.*, 2010; Ogan-Bekiroglu and Eskin, 2012; von Aufschnaiter *et al.*, 2008). What there is agreement on, however, is that concept knowledge and argumentation are not equally enhanced in interventions, but argumentation can lead to an activation of prior knowledge (Ogan-Bekiroglu and Eskin, 2012). Argumentation and prior knowledge have kind of a reciprocal relationship, as prior knowledge is necessary to build arguments meaningfully in the first place (de Lima Tavares *et al.*, 2010; Faize *et al.*, 2017; von Aufschnaiter *et al.*, 2008), but argumentation on the other hand can also help to develop an improved understanding of contexts, thereby relinking prior knowledge (de Lima Tavares *et al.*, 2010; Faize *et al.*, 2017). In a study by Cross *et al.* (2008) with high school students, it was also revealed that arguing and discussing content about which students have prior knowledge provides a sense of safety and competence. To address the inconsistencies in the relationship between concept knowledge and argumentation, the two scoring systems were rigorous in neglecting the use of concept knowledge when scoring the argument components and neglecting the structure of the argument built when scoring concept knowledge. Furthermore, the focus of the adaptive scaffold was not on the acquisition of new concept knowledge but on the activation of prior knowledge. Therefore, it is generally agreed that arguments cannot be appropriately build without concept knowledge. It was thus ensured in consultation with the instructor that the students have the necessary prior knowledge to be able to solve the tasks. The question is therefore not whether students have the necessary concept knowledge but whether they are able to activate it in the respective situation. Thus, the separation of the support into an argument and a concept knowledge part is initially legitime. However, to be able to verify an appropriate separation, it is to ask whether the adapted scaffolding groups improved their performance in the respective area of support.

Does the adaptive scaffold support students in improving their performance?

It was revealed that the performance gap between students was closed by the adaptive scaffold in terms of both the argumentation score and the concept knowledge score. However, it has remained an open question whether students in different groups merely converged or whether they were also able to improve within their group. To be able to determine the effectiveness of the scaffold, Wilcoxon signed-rank tests with Bonferroni-adjusted *p*-values were performed. The correlation coefficient *r* was defined as $.10 \leq r \leq .30$ as small effect, $.30 \leq r \leq .50$ as medium, and $r \geq .50$ as large (Cohen, 1992). The results are summarized graphically in Figure 9 and the respective values can be found in chapter 4.

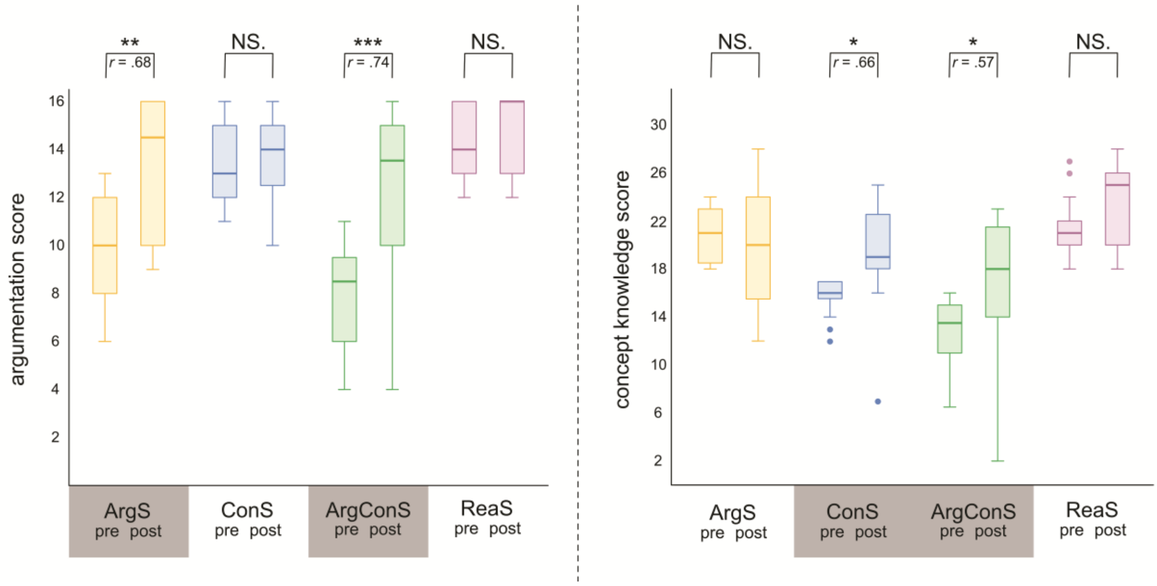


Fig. 9: Pre- and post- comparisons of the argumentation score and concept knowledge score of the four adapted scaffolding groups in the pre-measure and post-measure. Groups who received an adapted scaffold in the respective area of support are highlighted with a gray background. Horizontal stripes in box plots indicate median-values. Significance levels of the pre- and post-comparisons are indicated (NS. $p > .50$; * $p < .05$; ** $p < .01$; *** $p < .001$).

In terms of the argumentation score, the two groups that received support in structuring their arguments improved significantly with large effects (ArgS group $r = .68$, ArgConS group $r = .74$). The ConS and ReaS groups, neither of which received additional support in structuring their arguments, showed no significant changes when comparing pre- and post-measure. Comparable results are also found with respect to the concept knowledge score. The two groups that received support in applying concept knowledge improved significantly with large effects (ConS group $r = .66$, ArgConS group $r = .57$). However, the two groups without additional support in concept knowledge (ArgS group and ReaS group) did not show significant changes. Qualitative improvements were also evident. Mike, a student of the ArgConS group, serves as an example of this because the ArgConS group improved significantly in both the argumentation and concept knowledge scores. A built argument of Mike's pre-measure has already been shown in Figure 7. Figure 10 contrasts the built argument from the pre-measure with a built argument from the post-measure.

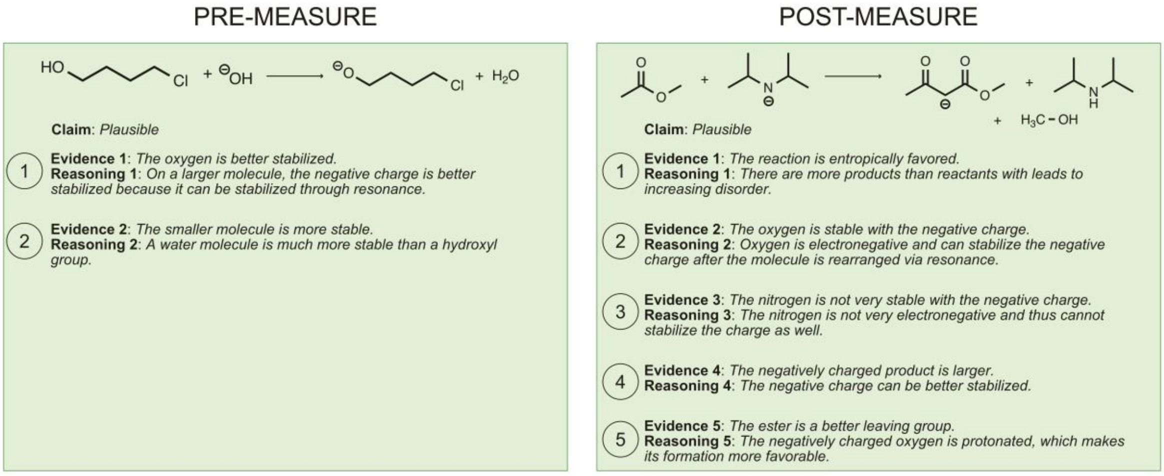


Fig. 10: Comparison of exemplary arguments built by Mike from the ArgConS group in the pre-measure (left) and the post-measure (right).

In comparing the two arguments, it becomes apparent that Mike improves as he formed more pieces of evidence and reasoning statements. Moreover, the justification was also based on the use of scientific principles. In applying concept knowledge, he still made some technical errors, but he made an effort to justify each piece of evidence with chemical concepts. Additionally, he used a variety of different chemical concepts in this process, which is often presented in the literature as an indication of a more intensive examination of the problem situation (Kraft *et al.*, 2010; Sevian and Talanquer, 2014; Talanquer, 2006; Weinrich and Talanquer, 2015). The link between concept knowledge and argumentation is evident in the post-measure in Mike's and the most other students' arguments, which is a necessary condition for the quality of an argument (Choi *et al.*, 2013; Sandoval and Millwood, 2005; von Aufschnaiter *et al.*, 2008). Thus, it can be assumed through analysis that the adaptive scaffold improved students' performance regarding their individual needs.

1.5 Conclusions

This dissertation demonstrates that argumentation on alternative reaction pathways can be used to diagnose and foster students in building concept knowledge and applying concept knowledge. Both diagnosis and fostering focus on the scientific practice of argumentation, which needs to be learned just like facts and concepts (Cooper, 2015; Crooks, 1988; Songer and Gotwals, 2012), and on linking argumentation and concept knowledge, which is applied in justifying claims with evidence and reasoning (Bricker and Bell, 2008; Driver *et al.*, 2000; Songer and Gotwals, 2012; von Aufschnaiter *et al.*, 2008). Argumentation on alternative reaction pathways also requires learners to consider the interplay of multiple chemical concepts and the influence that these concepts have on the reaction pathway. A simplified version of Toulmin's argumentation model was used as the basis for diagnosing and fostering argumentation: the *claim-evidence-reasoning* model (see Figure 1), which includes the core components of an argument (McNeill and Krajcik, 2012; McNeill *et al.*, 2006; Toulmin, 2003) but in an appropriate manner regarding the student cohort (Erduran, 2007; Lazarou and Erduran, 2020).

This dissertation shows new insights into how students build arguments on alternative reaction pathways and the challenges they experience. Addressing this question served as a diagnosis for tailoring a support for students. The diagnosis revealed that students could be assigned to three argumentation approaches based on their justification of claims through evidence and reasoning. This demonstrated that the number of justifications can improve the quality of an argument as students weighed multiple influencing factors. The analysis of justification further examined how students used reasoning in relation to chemical concepts. It was found that most students already use different chemical concepts, with the majority of reasoning statements based on electronic aspects. However, several students also experienced challenges in building justification and relied more on personal experiences, rather than using scientific principles as the basis of their reasoning. In addition to analyzing how students justify their claims, we also examined how students changed their initial claims after building arguments on alternative reaction pathways. This revealed that students in both studies often changed their initial claim after building arguments and more frequently predicted the correct reaction product as the most plausible product of the reaction. However, differences in the extent of this change emerged, as students from study 1 who built their arguments orally experienced a claim change more frequently than students from study 2 who built their arguments in writing, presumably due to a lack of activation of prior knowledge and a focus on memorized mechanisms. This was confirmed in the evaluation, as many students indicated that they initially answered tasks with memorized knowledge. However, in the reflection, the students also stated that they found the building of arguments on alternative reaction pathways valuable because they dealt with the tasks more intensively and also thought about the chemical processes running in the background. In the diagnosis it could therefore be determined that the students experienced challenges in the justification of their claim and in the use of concept knowledge as well as in the activation of their prior knowledge.

Based on the results of the diagnosis and the relevant literature, an adaptive scaffold was designed and tested to support students with regard to the structure of arguments and the application of concept knowledge. In doing so, the dissertation provides new insights into how an adaptive scaffold can support students regarding their individual needs. It was shown that students experienced expected challenges (1) in the structure of their arguments, (2) in applying concept knowledge, (3) in the structure of their arguments and in applying concept knowledge, or (4) in the formation of multivariable arguments. Moreover, the classification of students based on an argumentation score and concept knowledge score could be verified by quantitative analyses. Furthermore, the adaptive scaffold was able to close the performance gap that was found within the different scaffolding groups at the beginning. However, not only did the scaffolding groups converge, but the adaptive scaffold was also able to support the students in the area they experienced the greatest challenges, suggesting that the improvement was not due to a simple training effect.

The results, which both confirm the findings of the literature and provide many new insights, are promising and provide a rich foundation for future research to not only teach students the scientific practice of argumentation, but also to provide them with purposeful support in applying scientific principles in meaningful ways.

1.6 Outlook and Suggestions for Future Research

The design, testing, and evaluation of the dissertation project presented has shown that there is great potential to diagnose and foster students in building arguments on alternative reaction pathways in organic chemistry. In particular, the adaptive scaffold can be used in a variety of ways, as one can specifically emphasize individual aspects in classroom situations, such as structuring arguments or applying concept knowledge. Moreover, depending on the objectives, scoring can be neglected, as it is the most time-consuming component of the adaptive scaffold. Particularly in very large learning groups like in the U.S., where up to 600 students attend one lecture, manual scoring by the instructor is hardly feasible. However, to support students based on their individual needs, a diagnosis of strengths and challenges and thus scoring of students' answers is essential. One way to address this limitation is to design a computer-based adaptive learning system. This learning system can consist of two parts: (1) an automated scoring system to provide students with a timely and resource-efficient way to receive feedback, and based on this, (2) individualized support that is tailored to the strengths and challenges of the individual student. The following section presents approaches to how the development of a computer-based adaptive learning system might be implemented.

1.6.1 Extension for the diagnosis – Automatic scoring system

The goal of an automated scoring system is to be able to analyze and evaluate student answers with the help of a trained algorithm. In machine learning, a statistical model is created on the basis of an algorithm. This statistical model is then used as a decision model (Mitchell, 1997). The more heterogenous data are used to train the algorithm, the higher is the accuracy of the automated scoring system.

The idea of using an automated scoring system based on machine learning to analyze student answers is not new (Zhai *et al.*, 2020), but has already been implemented by several researchers in the natural sciences such as chemistry (Dood *et al.*, 2020; Dood *et al.*, 2018; Noyes *et al.*, 2020), biology (Anderson *et al.*, 2018; Haudek *et al.*, 2012), or physics (Mason and Just, 2016; Nakamura *et al.*, 2016). However, the majority of these automated scoring systems are item-specific, meaning that the scoring system can only be used for answers obtained with the exact same item. Yik *et al.* (2021) extended their scoring system by using 15 different items on the topic of Lewis acid-base model to train the machine to allow item-dependent scoring and thus a more versatile automated scoring system. Therefore, starting from the statistical model from Yik *et al.* (2021), an adapted version for the automated scoring system of the adaptive scaffold is designed. The goal is to perform both a structural analysis of student answers (*i.e.*, evidence and reasoning) and a content analysis (*i.e.*, the use of chemical concepts).

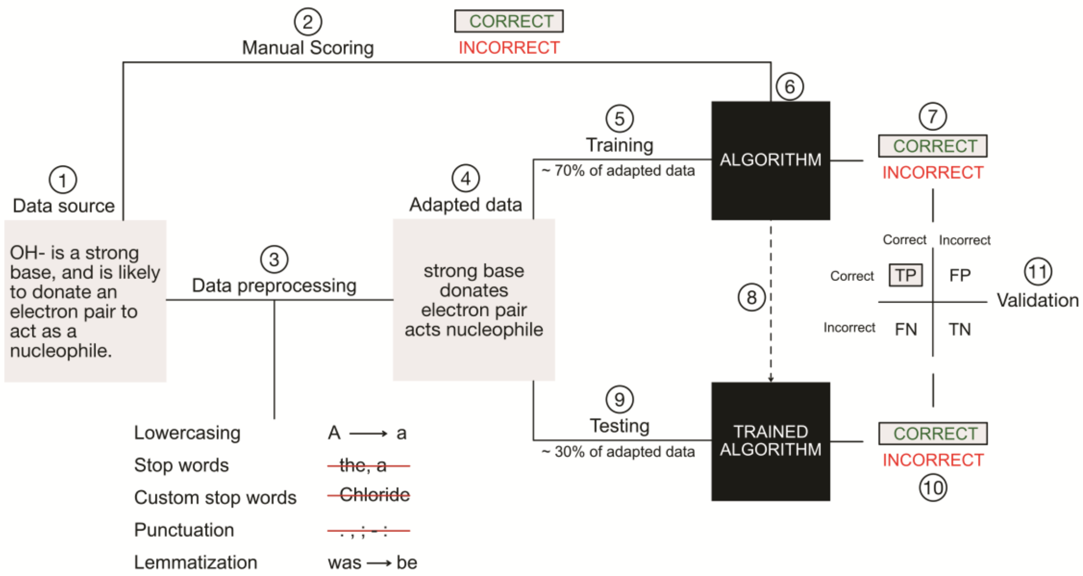


Fig. 11: Description of the process and evaluation of an automated scoring system using a machine learning approach. The four results of the validation are TP = true positive, FP = false positive, FN = false negative, TN = true negative.

Figure 11 illustrates the process for generating the automated scoring system. In the first step, the data source is created from the collected data from study 2. Here, the data is divided into individual units of meaning. The units of meaning are the individual fields of the evidence and reasoning table when building autonomous arguments on the alternative reaction pathways, resulting in a data size of about 2500 data sources. In the second step, these individual data sources must first be manually scored. For the development of the algorithm it is important that for each data source the corresponding analysis is performed (*i.e.*, correct or incorrect). After all data sources have been manually scored, the third step is the preprocessing of the data source for the generation of the decision model. The goal of the data preprocessing is to be able to analyze the data as generally as possible, so that the algorithm can later analyze a large number of student statements in a meaningful way. In the preprocessing, all upper-case letters are converted to lower case; stop words, which usually have a low information content (*e.g.*, the, a, …), custom stop words, such as names of reactants or products, as well as punctuation are removed. Finally, lemmatization takes place, where nouns, adjectives, adverbs, and conjugated verbs, are unified to a chosen form. For example, 'going' and 'went' become 'go'. This preprocessing turns the data source into an adapted data set (see the comparison of step 1 and step 4 in Figure 11). The adapted data is split, with about 70% of the data used to train the algorithm and the remaining 30% used to test the trained algorithm. When training the algorithm, the adapted data is combined with the manual analysis to train the model. As a result, each adapted data source receives a fixed analysis, for example, correct. Thus, the algorithm becomes a trained algorithm. To test the trained algorithm, the remaining 30% of the adapted data is used, which was not used for training before. As a result, the trained algorithm analyzes the data, *i.e.*, whether the student answers are analyzes as correct or incorrect. In the last step, a validation of the evaluation process takes place, comparing the agreement of the human-classified and computer-classified scores. The validation indicates the amount of percentage of the evaluation that matched (true positive and true negative) or mismatched (false positive or false negative). The higher the percentage of agreement, the better the algorithm can predict an evaluation of the data.

Already published results of several research groups were promising. Nevertheless, many people are skeptical about the use of algorithms in the evaluation of data, which is why certain limitations will be discussed here. First, to increase the significance of the algorithm, a large number of data should be used to train the algorithm (Zhou, 2016) to increase the accuracy of the automated scoring system. In addition, when training the algorithm, data from as many different universities and institutes as possible should be used to address different ways of building arguments and prior knowledge (Yik *et al.*, 2021). A common criticism is the accuracy of the algorithm compared to the manual evaluation of a human being. Here, it should be emphasized that both the algorithm and the human make mistakes, making neither alternative the perfect solution. Nevertheless, both manual and automated scoring achieve percentage accuracies between 84%-92% (Dood *et al.*, 2020; Dood *et al.*, 2018; Noyes *et al.*, 2020; Yik *et al.*, 2021; Zhai *et al.*, 2020). It should be noted that the automated scoring system is dependent from the human who designed and tested it because the human chooses the data which is used to train and test the automated scoring system. However, in automated scoring of the data, the length of student answers can be challenging. Whether an answer is scored as correct or incorrect is ultimately based on the combination of certain words. The longer an answer is, the more likely it is that the answer will be evaluated as false positive. The shorter the answer, the more likely it is that the answer will be rated false negative (Yik *et al.*, 2021).

However, the use of automated scoring systems offers great potential not only to save significant time for instructors, but also to allow the adaptive scaffold to be used in a way that is detached from the instructor, as the expert is represented by the computer to score the answers.

1.6.2 Extension of the fostering – Individualized adaptive scaffold

Based on a well-founded and repeatedly evaluated automated scoring system, it is also possible to adjust the adaptive scaffold more individually to the strengths and challenges of the students. Figure 12 provides a broad overview of how such a tool could be implemented. Comparable to the implementation of study 2, students equally work on the diagnostic scaffold. Scoring is then performed using the automated scoring system described in the previous section. Based on the scoring results, individual challenges can be diagnosed, which can refer to the structural level of an argument (*i.e.*, evidence and reasoning) as well as to the content level (*i.e.*, chemical concepts such as nucleophilicity and electrophilicity). Students receive individualized feedback in the process, which provides them with an overview of their strengths and challenges. Building on this, the adaptive scaffold contains different building blocks to address individual challenges in a meaningful way. Based on insights into effective learning and learning acquisition, students are provided with opportunities to meet their challenges, for example, through learning videos that connect explanations and representations by using dynamic highlighting techniques (Rodemer *et al.*, 2020; Rodemer *et al.*, 2022), eye movement modeling examples that highlight the gaze behavior of experts while they execute problem solving (Xie *et al.*, 2021), or question-answer tasks that are oriented, for example, to the science writing heuristic by supporting students in structuring the thought process and gradually achieving a solution (Hand *et al.*, 2002; Keys *et al.*, 1999)

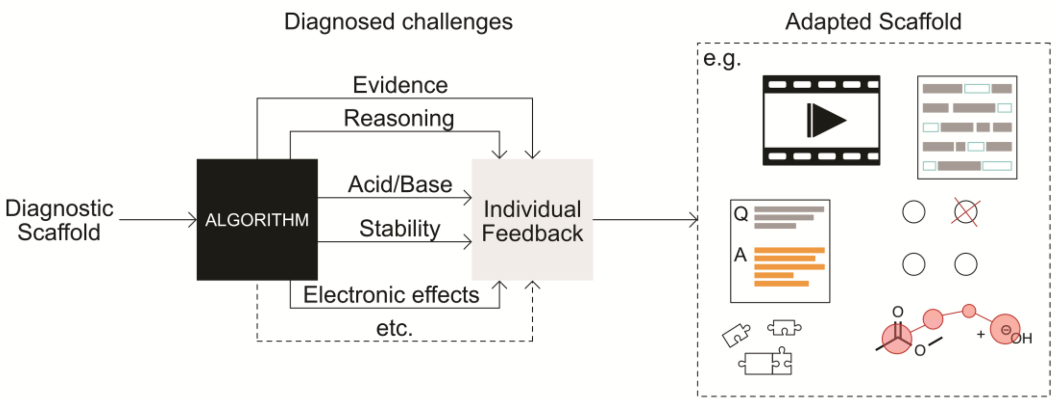


Fig. 12: Illustration of a feasible implementation of the automated scoring system, which is embedded in an adaptive learning system. Based on the challenges diagnosed in the automated scoring system, students receive individual feedback and different task formats for practicing and applying the respective content, which is part of an adapted scaffold.

In the implementation, for example, an interface in the form of a web page can be created on which different reactions are provided that have targeted foci. For example, reactions that focus on the competition between acid-base reactions and nucleophilic and electrophilic attacks or electronic effects on electrophilic aromatic substitutions. By providing a variety of reactions and their corresponding foci, instructors have the opportunity to use the adaptive scaffold to practice and reinforce after specific units in the curriculum, as well as to prepare for exams. By making the website available, students also have the opportunity to independently obtain feedback on their current learning status.

The two ideas presented for an extended use of diagnosing and fostering students to build arguments on alternative reaction pathways in organic chemistry represent only two exemplary ways in which the findings of this dissertation can impact future research.

1.7 References

Acar O., (2008), *Argumentation skills and conceptual knowledge of undergraduate students in a physics by inquiry class*, The Ohio State University.

Anderson C. W., de los Santos E. X., Bodbyl S., Covitt B. A., Edwards K. D., Hancock J. B., Lin Q., Morrison Thomas C., Penuel W. R. and Welch M. M., (2018), Designing educational systems to support enactment of the Next Generation Science Standards, *J. Res. Sci. Teach.*, **55** (7), 1026-1052.

Anzovino M. E. and Bretz S. L., (2015), Organic chemistry students' ideas about nucleophiles and electrophiles: the role of charges and mechanisms, *Chem. Educ. Res. Pract.*, **16** (4), 797-810.

Azevedo R., Cromley J. G. and Seibert D., (2004), Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia?, *Contemp. Educ. Psychol.*, **29** (3), 344-370.

Azevedo R., Moos D. C., Greene J. A., Winters F. I. and Cromley J. G., (2008), Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia?, *Educ. Technol. Res. Dev.*, **56** (1), 45-72.

Becker N., Rasmussen C., Sweeney G., Wawro M., Towns M. and Cole R., (2013), Reasoning using particulate nature of matter: An example of a sociochemical norm in a university-level physical chemistry class, *Chem. Educ. Res. Pract.*, **14** (1), 81-94.

Bell P. and Linn M. C., (2000), Scientific arguments as learning artifacts: Designing for learning from the web with KIE, *Int. J. Sci. Educ.*, **22** (8), 797-817.

Bernard H. R. and Bernard H. R., (2013), *Social research methods: Qualitative and quantitative approaches*, Sage.

Bhattacharyya G. and Bodner G. M., (2005), "It gets me to the product": How students propose organic mechanisms, *J. Chem. Educ.*, **82** (9), 1402-1407.

Bloom B., Engelhart M., Furst E., Hill W. and Krathwohl D., (1956), *Taxonomy of educational objectives*, New York: David McKay Company, Inc.

Bodé N. E., Deng J. M. and Flynn A. B., (2019), Getting Past the Rules and to the WHY: Causal Mechanistic Arguments When Judging the Plausibility of Organic Reaction Mechanisms, *J. Chem. Educ.*, **96**, 1068-1082.

Bodner G. M. and Domin D. S., (2000), Mental models: The role of representations in problem solving in chemistry, *U. Chem. Ed.*, **4** (1).

Brewer W. F., Chinn C. A. and Samarapungavan A., (1998), Explanation in scientists and children, *Minds Mach.*, **8** (1), 119-136.

Bricker L. A. and Bell P., (2008), Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education, *Sci. Educ.*, **92** (3), 473-498.

Cartrette D. P. and Bodner G. M., (2009), Non-mathematical problem solving in organic chemistry, *J. Res. Sci. Teach.*, **47** (6), 643-660.

Caspari I. and Graulich N., (2019), Scaffolding the structure of organic chemistry students' multivariate comparative mechanistic reasoning, *Int. J. Phys. Chem. Educ.*, **11** (2), 31-43.

Caspari I., Kranz D. and Graulich N., (2018), Resolving the complexity of organic chemistry students' reasoning through the lens of a mechanistic framework, *Chem. Educ. Res. Pract.*, **19** (4), 1117-1141.

Chang S. N. and Chiu M. H., (2008), Lakatos' scientific research programmes as a framework for analysing informal argumentation about socio-scientific issues, *Int. J. Sci. Educ.*, **30** (13), 1753-1773.

Chen C. H., (2014), An adaptive scaffolding e-learning system for middle school students' physics learning, *Australas. J. Educ. Technol.*, **30** (3), 342-355.

Choi A., Hand B. and Greenbowe T., (2013), Students' Written Arguments in General Chemistry Laboratory Investigations, *Res. Sci. Educ.*, **43** (5), 1763-1783.

Cohen J., (1992), A power primer, *Psychol. Bull.*, **112** (1), 155-159.

Cooper M. M., (2015), Why ask why?, *J. Chem. Educ.*, **92** (8), 1273-1279.

Cooper M. M., Corley L. M. and Underwood S. M., (2013), An investigation of college chemistry students' understanding of structure-property relationships, *J. Res. Sci. Teach.*, **50** (6), 699-721.
Cooper M. M., Kouyourndjian H. and Underwood S. M., (2016), Investigating Students' Reasoning about Acid-Base Reactions, *J. Chem. Educ.*, **93** (10), 1703-1712.
Cooper M. M. and Stowe R. L., (2018), Chemistry education research—From personal empiricism to evidence, theory, and informed practice, *Chem. Rev.*, **118** (12), 6053-6087.
Crooks T. J., (1988), The impact of classroom evaluation practices on students, *Rev. Educ. Res.*, **58** (4), 438-481.
Cross D., Taasoobshirazi G., Hendricks S. and Hickey D. T., (2008), Argumentation: A strategy for improving achievement and revealing scientific identities, *Int. J. Sci. Educ.*, **30** (6), 837-861.
Cruz-Ramirez de Arellano D. and Towns M. H., (2014), Students' understanding of alkyl halide reactions in undergraduate organic chemistry, *Chem. Educ. Res. Pract.*, **15** (4), 501-515.
de Lima Tavares M., Jiménez-Aleixandre M.-P. and Mortimer E. F., (2010), Articulation of conceptual knowledge and argumentation practices by high school students in evolution problems, *Sci. & Educ.*, **19** (6), 573-598.
Deng J. M. and Flynn A. B., (2021), Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items, *Chem. Educ. Res. Pract.*, **22** (3), 749-771.
Dood A. J., Dood J. C., Cruz-Ramirez de Arellano D., Fields K. B. and Raker J. R., (2020), Analyzing explanations of substitution reactions using lexical analysis and logistic regression techniques, *Chem. Educ. Res. Pract.*, **21** (1), 267-286.
Dood A. J., Fields K. B. and Raker J. R., (2018), Using lexical analysis to predict Lewis acid–base model use in responses to an acid–base proton-transfer reaction, *J. Chem. Educ.*, **95** (8), 1267-1275.
Driver R., Newton P. and Osborne J., (2000), Establishing the norms of scientific argumentation in classrooms, *Sci. Educ.*, **84** (3), 287-312.
Erduran S., (2007), in *Argumentation in Science Education*, Springer, pp. 47-69.
Faize F. A., Husain W. and Nisar F., (2017), A critical review of scientific argumentation in science education, *Eurasia J. Math. Sci.*, **14** (1), 475-483.
Field A., Miles J. and Field Z., (2012), *Discovering statistics using R*, Sage publications.
Flynn A. B., (2011), Developing Problem-Solving Skills through Retrosynthetic Analysis and Clickers in Organic Chemistry, *J. Chem. Educ.*, **88** (11), 1496-1500.
Ford M., (2008), Disciplinary authority and accountability in scientific practice and learning, *Sci. Educ.*, **92** (3), 404-423.
Gigerenzer G. and Gaissmaier W., (2011), Heuristic decision making, *Annu. Rev. Psychol.*, **62**, 451-482.
Govier T., (1987), *Problems in Argument Analysis and Evaluation* Dordrecht: De Gruyter Mouton.
Graulich N., (2014), Intuitive judgments govern students' answering patterns in multiple-choice exercises in organic chemistry, *J. Chem. Educ.*, **92** (2), 205-211.
Greene B. A. and Land S. M., (2000), A qualitative analysis of scaffolding use in a resource-based learning environment involving the World Wide Web, *J. Educ. Comput. Res.*, **23** (2), 151-179.
Grove N. P., Cooper M. M. and Cox E. L., (2012), Does mechanistic thinking improve student success in organic chemistry?, *J. Chem. Educ.*, **89** (7), 850-853.
Hand B., Prain V. and Wallace C., (2002), Influences of writing tasks on students' answers to recall and higher-level test questions, *Res. Sci. Educ.*, **32** (1), 19-34.
Haudek K. C., Prevost L. B., Moscarella R. A., Merrill J. and Urban-Lurain M., (2012), What are they thinking? Automated analysis of student writing about acid–base chemistry in introductory biology, *CBE Life Sci. Educ.*, **11** (3), 283-293.
Henderson J. B. and Osborne J. F., (2019), in *Argumentation in Chemistry Education*, Royal Society of Chemistry, pp. 79-105.
Hogan K. and Maglienti M., (2001), Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions, *J. Res. Sci. Teach.*, **38** (6), 663-687.

Hosbein K. N., Lower M. A. and Walker J. P., (2021), Tracking Student Argumentation Skills across General Chemistry through Argument-Driven Inquiry Using the Assessment of Scientific Argumentation in the Classroom Observation Protocol, *J. Chem. Educ.*, **98** (6), 1875-1887.

Jiménez-Aleixandre M. P. and Erduran S., (2007), in *Argumentation in Science Education*, Springer, pp. 3-27.

Kang H., Thompson J. and Windschitl M., (2014), Creating Opportunities for Students to Show What They Know: The Role of Scaffolding in Assessment Tasks, *Sci. Educ.*, **98** (4), 674-704.

Keys C. W., Hand B., Prain V. and Collins S., (1999), Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science, *J. Res. Sci. Teach.*, **36** (10), 1065-1084.

Kim N. J., Vicentini C. R. and Belland B. R., (2022), Influence of scaffolding on information literacy and argumentation skills in virtual field trips and problem-based learning for scientific problem solving, *Int. J. Sci. Math. Educ.*, **20** (2), 215-236.

Kraft A., Strickland A. M. and Bhattacharyya G., (2010), Reasonable reasoning: multi-variate problem-solving in organic chemistry, *Chem. Educ. Res. Pract.*, **11** (4), 281-292.

Kranz D., Schween M. and Graulich N., (2022), Patterns of reasoning - exploring the interplay of students' work with a scaffold and their conceptual knowledge in organic chemistry, *Chem. Educ. Res. Pract.*, submitted manuscript.

Kuhn D., (1992), Thinking as argument, *Harv. Educ. Rev.*, **62** (2), 155-179.

Kuhn L. and Reiser B., Structuring activities to foster argumentative discourse, 2006.

Lajoie S. P., (2005), Extending the scaffolding metaphor, *Instr. Sci.*, **33** (5-6), 541-557.

Lazarou D. and Erduran S., (2020), "Evaluate What I Was Taught, Not What You Expected Me to Know": Evaluating Students' Arguments Based on Science Teachers' Adaptations to Toulmin's Argument Pattern, *J. Sci. Teach. Educ.*, **32** (3), 306-324.

Lin T.-C., Hsu Y.-S., Lin S.-S., Changlai M.-L., Yang K.-Y. and Lai T.-L., (2012), A review of empirical evidence on scaffolding for science education, *Int. J. Sci. Math. Educ.*, **10** (2), 437-455.

Lombardi D., Brandt C. B., Bickel E. S. and Burg C., (2016), Students' evaluations about climate change, *Int. J. Sci. Educ.*, **38** (8), 1392-1414.

Longino H. E., (1990), *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton University Press.

Luo X. L., Wei B., Shi M. and Xiao X., (2020), Exploring the impact of the reasoning flow scaffold (RFS) on students' scientific argumentation: based on the structure of observed learning outcomes (SOLO) taxonomy, *Chem. Educ. Res. Pract.*, **21** (4), 1083-1094.

Mangiafico S., (2021), rcompanion: Functions to Support Extension Education Program Evaluation.

Mason R. A. and Just M. A., (2016), Neural representations of physics concepts, *Psychol. Sci.*, **27** (6), 904-913.

McDonald C. V., (2010), The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science, *J. Res. Sci. Teach.*, **47** (9), 1137-1164.

McNeill K. L. and Krajcik J., (2012), *Book study facilitator's guide: Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*, New York: Pearson Allyn & Bacon.

McNeill K. L., Lizotte D. J., Krajcik J. and Marx R. W., (2006), Supporting students' construction of scientific explanations by fading scaffolds in instructional materials, *J. Learn. Sci.*, **15** (2), 153-191.

Mitchell T. M., (1997), *Machine learning*McGraw-Hill New York.

Momsen J. L., Long T. M., Wyse S. A. and Ebert-May D., (2010), Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills, *CBE Life Sci. Educ.*, **9** (4), 435-440.

Moon A., Stanford C., Cole R. and Towns M., (2016), The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry, *Chem. Educ. Res. Pract.*, **17** (2), 353-364.

Morewedge C. K. and Kahneman D., (2010), Associative processes in intuitive judgment, *Trends Cogn. Sci.*, **14** (10), 435-440.
Nakamura C. M., Murphy S. K., Christel M. G., Stevens S. M. and Zollman D. A., (2016), Automated analysis of short responses in an interactive synthetic tutoring system for introductory physics, *Phys. Rev. Phys. Educ. Res.*, **12** (1), 010122.
Noroozi O., Kirschner P. A., Biemans H. J. A. and Mulder M., (2017), Promoting Argumentation Competence: Extending from First- to Second-Order Scaffolding Through Adaptive Fading, *Educ. Psychol.*, **30** (1), 153-176.
Novak A. M. and Treagust D. F., (2018), Adjusting claims as new evidence emerges: Do students incorporate new evidence into their scientific explanations?, *J. Res. Sci. Teach.*, **55** (4), 526-549.
Noyes K., McKay R. L., Neumann M., Haudek K. C. and Cooper M. M., (2020), Developing computer resources to automate analysis of students' explanations of London dispersion forces, *J. Chem. Educ.*, **97** (11), 3923-3936.
Ogan-Bekiroglu F. and Eskin H., (2012), Examination of the relationship between engagement in scientific argumentation and conceptual knowledge, *Int. J. Sci. Math. Educ.*, **10** (6), 1415-1443.
Ohlsson S., (2002), *Generating and understanding qualitative explanations*, New York: Lawrence Erlbaum Associates
Osborne J. F. and Patterson A., (2011), Scientific argument and explanation: A necessary distinction?, *Sci. Educ.*, **95** (4), 627-638.
Pabuccu A. and Erduran S., (2017), Beyond rote learning in organic chemistry: the infusion and impact of argumentation in tertiary education, *Int. J. Sci. Educ.*, **39** (9), 1154-1172.
Patrick W., (1924), What kind of research is essential to good teaching?, *J. Chem. Educ.*, **1** (1), 16.
Pea R. D., (2004), The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity, *J. Learn. Sci.*, **13** (3), 423-451.
Petritis S. J., Kelley C. and Talanquer V., (2021), Exploring the impact of the framing of a laboratory experiment on the nature of student argumentation, *Chem. Educ. Res. Pract.*, **22** (1), 105-121.
Petritis S. J., Kelley C. and Talanquer V., (2022), Analysis of factors that affect the nature and quality of student laboratory argumentation, *Chem. Educ. Res. Pract.*, **23**, 257-274.
Pickering A., (2010), in *The Mangle of Practice*, University of Chicago Press.
Popova M. and Bretz S. L., (2018), "It's Only the Major Product That We Care About in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams, *J. Chem. Educ.*, **95** (7), 1086-1093.
Puntambekar S. and Hubscher R., (2005), Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed?, *Educ. Psychol.*, **40** (1), 1-12.
R Core Team, (2021), *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing.
Revelle W. R., (2021), psych: Procedures for personality and psychological research.
Rodemer M., Eckhard J., Graulich N. and Bernholt S., (2020), Decoding Case Comparisons in Organic Chemistry: Eye-Tracking Students' Visual Behavior, *J. Chem. Educ.*, **97** (10), 3530-3539.
Rodemer M., Lindner M. A., Eckhard J., Graulich N. and Bernholt S., (2022), Dynamic signals in instructional videos support students to navigate through complex representations: An eye-tracking study, *Appl. Cogn. Psychol.*
Rosenthal R., (1991), *Meta-analytic procedures for social research*, Newbury Park, CA: Sage.
Sadler T. D., (2004), Informal reasoning regarding socioscientific issues: A critical review of research, *J. Res. Sci. Teach.*, **41** (5), 513-536.
Sadler T. D. and Zeidler D. L., (2004), The morality of socioscientific issues: Construal and resolution of genetic engineering dilemmas, *Sci. Educ.*, **88** (1), 4-27.
Sandoval W. A., Enyedy N., Redman E. H. and Xiao S., (2019), Organising a culture of argumentation in elementary science, *Int. J. Sci. Educ.*, **41** (13), 1848-1869.

Sandoval W. A. and Millwood K. A., (2005), The quality of students' use of evidence in written scientific explanations, *Cogn. Instr.*, **23** (1), 23-55.

Sevian H. and Talanquer V., (2014), Rethinking chemistry: a learning progression on chemical thinking, *Chem. Educ. Res. Pract.*, **15** (1), 10-23.

Songer N. B. and Gotwals A. W., (2012), Guiding explanation construction by children at the entry points of learning progressions, *J. Res. Sci. Teach.*, **49** (2), 141-165.

Stanford C., Moon A., Towns M. and Cole R., (2016), Analysis of instructor facilitation strategies and their influences on student argumentation: A case study of a process oriented guided inquiry learning physical chemistry classroom, *J. Chem. Educ.*, **93** (9), 1501-1513.

Stowe R. L. and Cooper M. M., (2019), Arguing from spectroscopic evidence, *J. Chem. Educ.*, **96** (10), 2072-2085.

Stowe R. L., Scharlott L. J., Ralph V. R., Becker N. M. and Cooper M. M., (2021), You Are What You Assess: The Case for Emphasizing Chemistry on Chemistry Assessments, *J. Chem. Educ.*, **98** (8), 2490-2495.

Sweller J., (2003), in *Psychology of Learning and Motivation: Advances in Research and Theory*, San Diego, USA: Elsevier Science, pp. 215-266.

Talanquer V., (2006), Commonsense chemistry: A model for understanding students' alternative conceptions, *J. Chem. Educ.*, **83** (5), 811-816.

Talanquer V., (2013), in *Concepts of matter in science education*, Springer, pp. 331-346.

Talanquer V., (2017), Concept Inventories: Predicting the Wrong Answer May Boost Performance, *J. Chem. Educ.*, **94** (12), 1805-1810.

Toulmin S. E., (2003), *The Uses of Argument*, Cambridge University Press.

Tüzün Ü. N., Tüysüz M. and Türk G. E., (2021), The Effect of Argumentation-Based Organic Chemistry Teaching on Students' Argument Construction Skills, *Int. J. Contemp. Educ. Res.*, **8** (2), 46-56.

Van de Pol J., Volman M. and Beishuizen J., (2010), Scaffolding in teacher–student interaction: A decade of research, *Educ. Psychol.*, **22** (3), 271-296.

von Aufschnaiter C., Erduran S., Osborne J. and Simon S., (2008), Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge, *J. Res. Sci. Teach.*, **45** (1), 101-131.

Vygotsky L. S., (1980), *Mind in society: The development of higher psychological processes*, Harvard University Press.

Walker J. P., Sampson V., Grooms J., Anderson B. and Zimmerman C. O., (2012), Argument-driven inquiry in undergraduate chemistry labs: The impact on students' conceptual understanding, argument skills, and attitudes toward science, *J. Coll. Sci. Teach.*, **41** (4), 74-81.

Walker J. P., Van Duzor A. G. and Lower M. A., (2019), Facilitating Argumentation in the Laboratory: The Challenges of Claim Change and Justification by Theory, *J. Chem. Educ.*, **96** (3), 435-444.

Walqui A., (2006), Scaffolding instruction for English language learners: A conceptual framework, *Int. J. Biling. Educ. Biling.*, **9** (2), 159-180.

Watts F. M., Schmidt-McCormack J. A., Wilhelm C. A., Karlin A., Sattar A., Thompson B. C., Gere A. R. and Shultz G. V., (2020), What students write about when students write about mechanisms: analysis of features present in students' written descriptions of an organic reaction mechanism, *Chem. Educ. Res. Pract.*, **21** (4), 1148-1172.

Watts F. M., Zaimi I., Kranz D., Graulich N. and Shultz G. V., (2021), Investigating students' reasoning over time for case comparisons of acyl transfer reaction mechanisms, *Chem. Educ. Res. Pract.*, **22** (2), 364-381.

Weinrich M. L. and Talanquer V., (2015), Mapping students' conceptual modes when thinking about chemical reactions used to make a desired product, *Chem. Educ. Res. Pract.*, **16** (3), 561-577.

Wickham H., (2016), *ggplot2: Elegant graphics for data analysis*, New York: Springer Verlag.

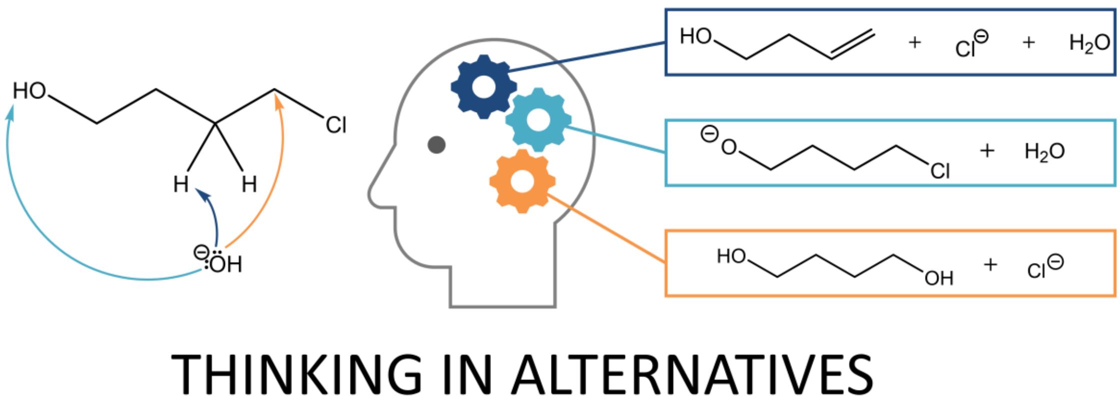
Wilson R. A. and Keil F., (1998), The shadows and shallows of explanation, *Minds Mach.*, **8** (1), 137-159.

Wood D., Bruner J. S. and Ross G., (1976), The role of tutoring in problem solving, *J. Child Psychol. Psychiatry*, **17** (2), 89-100.

Xie H., Zhao T., Deng S., Peng J., Wang F. and Zhou Z., (2021), Using eye movement modelling examples to guide visual attention and foster cognitive performance: A meta-analysis, *J. Comput. Assist. Learn.*, **37** (4), 1194-1206.
Yik B. J., Dood A. J., de Arellano D. C. R., Fields K. B. and Raker J. R., (2021), Development of a machine learning-based tool to evaluate correct Lewis acid-base model use in written responses to open-ended formative assessment items, *Chem. Educ. Res. Pract.*, **22** (4), 866-885.
Zhai X., Yin Y., Pellegrino J. W., Haudek K. C. and Shi L., (2020), Applying machine learning in science assessment: a systematic review, *Stud. Sci. Educ.*, **56** (1), 111-151.
Zhou Z. H., (2016), Learnware: on the future of machine learning, *Frontiers of Computer Science*, **10** (4), 589-590.

2 Thinking in Alternatives – A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry

In Chapter 2, the original paper is represented. It is reproduced from Lieber and Graulich (2020) with permission of the American Chemical Society.



Lieber, L. & Graulich, N. (2020). Thinking in Alternatives – A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry. *Journal of Chemical Education*, 97 (10), 3731-3738.

JOURNAL OF CHEMICAL EDUCATION

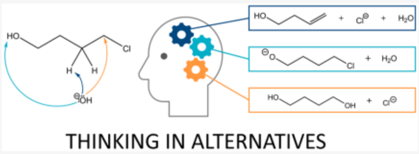
pubs.acs.org/jchemeduc Activity

Thinking in Alternatives—A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry

Leonie Lieber and Nicole Graulich*

Cite This: *J. Chem. Educ.* 2020, 97, 3731−3738 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: In organic chemistry, meaningful learning is essential when reflecting about multiple reaction pathways and selecting reaction centers—topics that require a complex and multivariate reasoning approach to problem-solving. Meaningful learning is characterized by abstract, analytical thinking that might be time-consuming, whereas rote learning, on the other side of the continuum, facilitates simple heuristic-based recall of information. To challenge organic chemistry students to shift their problem-solving approach from rote learning to meaningful learning, we designed task sequences that initially required students to use heuristics followed by an enforced reflection about alternative reaction pathways. Twenty-nine students who were enrolled in a third-year organic chemistry course participated in the study. They solved two tasks, each with four subtasks, which sequentially provoked the use of different problem-solving approaches. The tasks are centered on the chemical concept of nucleophilicity and electrophilicity. In the first step, students were prompted to predict the product of an easy and familiar reaction. This step aimed at provoking the use of heuristics. Afterward, students received five alternative product cards and were prompted to reflect on the plausibility and the alternative mechanistic pathways leading to these products. Evaluation of these task sequences revealed (a) to what extent students were engaged in reflecting on their own problem-solving approaches and (b) the aspects students perceived as important for adopting a more meaningful approach to solving typical tasks in organic chemistry. Implications and limitations of the designed tasks are discussed.

KEYWORDS: *Upper-Division Undergraduate, Organic Chemistry, Problem Solving/Decision Making, Acids/Bases, Mechanisms of Reactions, Nucleophilic Substitution*

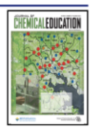
■ INTRODUCTION

Problem-solving is an essential ability possessed by successful experts in a given discipline because it allows them to achieve higher-order thinking.[1,2] Experts often approach problems differently than novices. They explore the problem, plan strategies, and group information, whereas novices tend to use commonplace solutions and to memorize given information.[3] Experts are efficient because they organize their knowledge in a way that relieves their working memory capacities.[4] Thus, they are able to analyze information in detail and restructure the problem because they have more experience in their professional domain.[5] With regard to organic chemistry, Cartrette and Bodner[5] wrote that successful problem-solving consists of "rationalizing trends in reactivity, determining spatial relationships, and devising coherent mechanisms to explain or predict chemical transformations" (p. 645). They also mentioned that more successful participants work in a more methodical manner as they always use the same strategy when solving a problem, no matter how simple or difficult the problem appears to be.[5] As students often struggle with problem-solving,[3] a key goal of chemistry education is to help students develop their problem-solving abilities.

Several studies have investigated problem-solving in organic chemistry. Bodé et al.[6] explored students' ability to construct arguments. Second-semester students received two reaction mechanisms of nucleophilic substitutions and had to compare these two reactions, predict the plausibility, and provide a reasoned explanation.[6] The researchers found that students' correct claims were supported by evidence but did not have the expected conceptual depth; students' incorrect claims were often based on incorrect evidence. The researchers concluded that even though comparing and contrasting reaction pathways seemed to engage students in problem-solving, the students did not use the highest mode of reasoning (e.g., multi-component mode of reasoning) in their problem-solving approaches.[6] Connecting contrasting cases with reflection prompts may direct students to a greater awareness of their problem-solving approaches.

DeCocq and Bhattacharyya[7] gave second-year undergraduate students elementary mechanistic steps to propose the product and electron-pushing mechanism of a reaction and

Received: March 11, 2020
Revised: August 21, 2020
Published: September 10, 2020



https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

later gave them the overall reaction, aiming to investigate whether external representations might evoke problem-solving strategies.[7] They stated that traditional tasks such as giving students reagents, solvents, and products may not be the most effective assessment to engage students in problem-solving. They concluded that asking participants to consider intermediates and multiple reaction centers might have helped them to think more deeply about chemical transformations.[7] Bhattacharyya and Bodner[8] investigated problem-solving in organic synthesis with electron-pushing tasks. They observed that "rather than solving chemical problems, [students] were essentially playing with puzzles" (p. 1406). We surmise that this was because students focused on remembering reaction products.

To summarize, a key goal of chemistry education research that focuses on problem-solving is as follows: To help students become aware that their analysis about reaction mechanisms requires them to consider and apply underlying chemical concepts and principles in order to make a judgment about the plausibility of reactions.

■ THEORETICAL PERSPECTIVE

On a continuum from meaningful learning to rote learning, it has been found that students mostly stand at the rote learning end.[9] In organic chemistry, meaningful learning is essential because organic chemistry is a process-oriented field in which complex, multivariate reasoning skills are required for effective problem-solving.[4,10] Many students have difficulty in predicting the outcome of reactions using strategies such as balancing potential mechanistic pathways and reflecting on underlying concepts; instead, they make use of product-oriented thinking and predict products without weighing chemical concepts nor considering mechanisms.[11] Under the lens of human decision-making, rote and meaningful learning can be defined as type 1 and type 2 reasoning.[12]

In type 1 reasoning, rote learning is intuitive and the decision-making process is quick. Rote learning techniques, such as heuristics, rely on patterns of recognition and familiarity.[13,14] When solving tasks in organic chemistry, students often use "one-reason" decision-making to explain changes or properties of reactions.[15] Students tend to take explicit features of representations into consideration rather than analyze the implicit properties or influencing variables.[16−20] Although heuristics are useful tools for quick decision-making, they can lead to alternative conceptions when fundamental knowledge is lacking.[21]

By using heuristics, students develop explanations that may be accepted, even if they do not understand its underlying meaning. Thus, an explanation can create an illusion of competence when the student is not prompted to justify the answer.[13,22] Rote learning strategies are often used when the cost of gaining information is high (e.g., in an information-rich organic chemistry course).[15,23] By neglecting to use appropriate meaningful techniques and relying instead on rote memorization, students can encounter problems.[9]

Type 2 reasoning is characterized by abstract and analytical thinking and is considered to be time-consuming.[12] Even though students tend to use simple heuristic strategies that minimize their cognitive efforts, slow, analytical reasoning may allow them to think through the underlying chemical processes and to predict unknown reactions and alternative pathways correctly and in depth. Therefore, students need to be aware that both, explicit and implicit features of a reaction, should to be included in the decision-making process. However, meaningful strategies, such as thinking about alternative mechanistic pathways, require the processing and application of various pieces of information in the working memory, which make it cognitively demanding and rather slow.[24]

Analytical thinking can be achieved by using moments of cognitive dissonance (e.g., encountering inconsistences) to recognize features that were previously irrelevant or overlooked as relevant.[25] Cognitive dissonance occurs when students perceive a contrast between their own knowledge and observations (e.g., of chemical phenomena or other peoples' beliefs).[26] This dissonance is an uncomfortable psychological state that presumably leads students to defend or revise their positions to reduce dissonance and reach consonance.[25,27] As a result, students reflect on their possibly erroneous approaches to problem-solving. In particular, reflecting on one's own incorrect approaches to problem-solving has a positive effect on knowledge acquisition.[28] An error reflection can expand a learner's cognitive model by adding to it the ability to discriminate between incorrect strategies and misconceptions. Thus, a more comprehensive cognitive model can be generated, fostering a more reflective approach to problem-solving.[28] Loibl and Leuders[29] found that the most beneficial effects of reflection can occur when students generate and then compare erroneous solutions with correct solutions, which can lead to a revision of their cognitive models.[29]

The literature review shows that students encounter difficulties in learning organic chemistry and tend to use rote learning as their preferred problem-solving approach. Therefore, we designed task sequences to challenge students' problem-solving approaches by creating moments of cognitive dissonance. The goal of our study was to investigate (a) how students rate the difficulty of these newly designed tasks, (b) if the task design encourages them to reflect on their own problem-solving approaches, and (c) to what extent students adopt a more meaningful approach when solving these tasks.

■ METHODS

Activity

The following activity, which included two tasks, was designed to challenge students' problem-solving approaches in organic chemistry and was quite unlike the questions and task design in students' traditional learning environment. The tasks were centered on nucleophilicity and electrophilicity, two prominent concepts that are susceptible to misconceptions in organic chemistry courses.[16,17,30,31] Both tasks (A/B) contained four subtasks (A1−A4/B1−B4). An overview of task A is shown in Figure 1; additional material is available in the Supporting Information. Since both tasks are based on the same principle, the design of the subtasks in task A is provided as an illustrative example.

The first subtask, A1, contained a typical, familiar reaction that is normally discussed in introductory organic chemistry courses. This task was designed to provoke the use of intuitive heuristics as students can easily recognize explicit surface features of an S_N2 substitution reaction (e.g., a good leaving group and a nucleophile).[20] For instance, students might remember S_N2 substitution reactions and solve the task by recalling the solution. Subtask A1 was purposely designed to guarantee easy success from rote learning. Subtask A2 appeared as easy as subtask A1 because the reaction differed

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

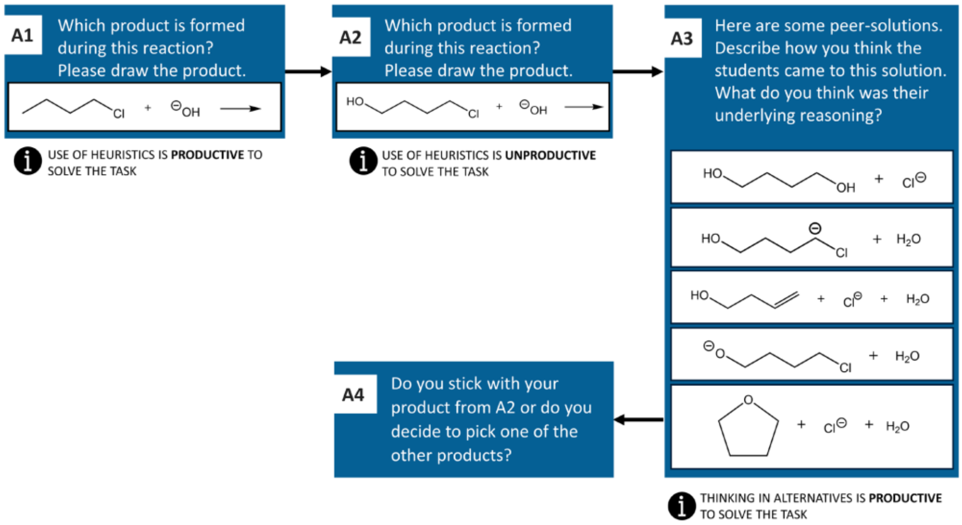


Figure 1. Prompts and product cards for task A. Task A and B each lasted each between 30 to 60 min.

only in one additional feature on the surface level (i.e., the hydroxy group). A2 was designed to provoke the use of the same problem-solving approach as A1, but this time a recall of the same S_N2 reaction pattern would not be productive. In the next subtask, A3, students were provided with product cards showing five alternative products for the reaction from subtask A2. This was a crucial step that could lead students to substitute intuitive approaches with analytical processes because, in this subtask, students would be compelled to take an outsiders' perspective.[32] These alternative reaction product cards proposed various solutions that were partly correct or plausible. The A3 step was designed such that cognitive dissonance could occur in students' reasoning process because they were confronted with observations or results that may contradict their existing beliefs.[26] As per previous research, cognitive dissonance causes students discomfort and leads them to defend or correct their decision,[27] thus initiating a more analytical thinking process in which they have to weigh chemical concepts to make a decision.[19] Lastly, subtask A4 provided students with the opportunity to defend or revise the generated product in A2 after being prompted to reflect on the given alternatives.

Implementation

The tasks were assigned to students enrolled in an Organic Chemistry 3 course at a German university at the beginning of the fall semester of 2019. Forty students, all chemistry majors, were enrolled in the course. Twenty-nine out of 40 students (they identified themselves as males (18) and females (11)) volunteered to participate in individual, think-aloud interviews guided by a semistructured protocol and conducted by the first author. Students were recruited at the beginning of the Organic Chemistry 3 course via an announcement in class. All students were native German speakers, so the interviews were conducted in German. The 29 interviews took place at the Institute of Chemistry Education during which students solved the tasks and answered questions on the Likert scale questionnaire items. All participants were provided with information about their rights and the manner of data collection and gave written permission for the data collection. They also gave permission to the authors to analyze and publish their data (i.e., reproduce transcripts and drawings in scientific publications). Institutional Review Board approval is not required at German universities, but the interviews were conducted as per the ethical guidelines of the German Data Privacy Act, the recommendation of the German Research Foundation, and the guidelines of the authors' university. The participants had the opportunity to terminate their interview at any point. During the interviews, the participants received two tasks (A/B) with four subtasks (A1–A4/B1–B4). Task A is shown in Figure 1. Additional material is available in the Supporting Information.

Data Collection and Data Analysis

The collected data included participants' demographic information, pseudonymized scans of participants' work sheets, and audio and video recordings. Verbal utterances were transcribed verbatim and data analysis was conducted in German. Quotes, selected for use in publications were translated into English. To guarantee validity of translations and meaning, quotes were checked multiple times by the research group and a German English student teacher.

To evaluate students' experiences with the designed tasks, quantitative and qualitative data analyses were conducted. Quantitative analysis was conducted on the Likert-scale data collected from students for each subtask. Specifically, each subtask was accompanied by a Likert scale item to rate the level of subjective difficulty, from 1 (easy) to 5 (difficult), each participant experienced in completing the subtask. After solving a subtask, all 29 participants immediately determined the rating of subjective difficulty of the subtask without knowing what the content of the next subtask would be. To

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

ascertain the subjective difficulty, we calculated the mean and standard deviation of students' rating for each subtask. Additionally, we compared students' rating on the difficulty level of each subtask with their subtask solutions.

Qualitative analysis was conducted on participants' interview data and participants responses to the evaluation questions posed at the end of task A and task B. The evaluation questions addressed different aspects. The participants were asked (a) how they had experienced the tasks in general, (b) how they had approached typical organic chemistry tasks before the interview and during the interview, (c) what aspects of the tasks they considered important and usable in their learning environments, and (d) what was their opinion on how these designed task sequences differed from typical, familiar tasks they encounter in their typical learning environment (e.g., give-the-product tasks). To evaluate students' experience with the tasks, students' transcripts were sorted based on the research questions. Representative quotes are represented as examples in the Results and Discussion section.

■ RESULTS AND DISCUSSION

How Do Students Rate the Difficulty of Solving the Tasks?

Two tasks were administered to 29 students from an Organic Chemistry course to test their problem-solving approach for organic chemistry tasks. To evaluate the subtasks, each subtask included a Likert scale item that participants used to rate the level of subjective difficulty of the tasks from 1 (easy) to 5 (difficult). All 29 participants responded to these items. These data are reported in Table 1 and indicate, as expected, that students perceived the tasks to be rather difficult. A1 was rated as the easiest subtask because the rote learning approach was successful. A2 differed from A1 in only one additional surface feature but one that students found more difficult. A3 provided five alternative products compared to A2. Students had to compare and weigh different chemical concepts in which the previously productive heuristics were not helpful. As a result, students rated subtask A3 more difficult than A1/A2. The same trend is noticeable for task B. Specifically, B1, as the initial subtask, was rated easier in relation to the more complex subtask B3. B2 seemed easier than B1 because many students recognized the reagent lithium diisopropylamide (LDA) as a base even though they could not explain why LDA reacts as a strong base but as a weak nucleophile.

Table 1. Results of the Likert Scale[a] After Completion of Each Subtask

Subtask	X[b]	SD[c]
A1	1.4	0.6
A2	2.2	1.1
A3	2.8	0.9
B1	2.5	0.9
B2	2.2	1.1
B3	2.9	0.9

[a]The scale for response ranged from 1 to 5: 1 = easy, 2 = rather easy, 3 = medium, 4 = rather difficult, 5 = difficult. [b]Means. [c]Standard deviation.

Compared with students' subjective rating of the difficulty level of the subtasks, students' proposed products were insightful. The total number of products was higher than the number of participants because participants had the opportunity to propose more than one plausible product. Subtasks A1 and B1 were designed to be easy and this easiness factor was supported by students' rating on the Likert scale and the proposed products. Twenty-six out of the 29 students proposed an alcohol via a nucleophilic substitution and six students proposed an alkene via an elimination for subtask A1. Both proposed products were correctly chosen. Twenty students proposed an amide via a nucleophilic acyl substitution for subtask B1. In subtask A2, 23 students proposed a diol as the product of the reaction which was incorrect, and only three students proposed the correct product, tetrahydrofuran (THF). After discussing several product cards, 11 students still proposed the diol as their main product for the reaction, whereas 16 proposed THF. In subtask B2, 13 students proposed an enolate via an acid–base reaction as the correct product because they recognized LDA as a strong base. This recognition effect can also be seen in students' rating of difficulty level as the participants rated subtask B2 easier than subtask B1. Nevertheless, only four students proposed an enolate as a precursor for an aldol reaction. After thinking through the given product cards, 21 students proposed the correct product of the aldol reaction as the main product. It became apparent that students rarely proposed the correct products of the given reaction already in subtask A2/B2. While discussing the product cards, students had to weigh chemical concepts that often led to a correction of their proposed product and an increasing number of correctly proposed products.

To evaluate the tasks qualitatively and to what extent students reflected about their problem-solving approaches, all participants were asked the same questions after task A and B. Representative quotes are shown in the relevant discussion below.

Does the Task Design Initiate Students' Reflection about Their Problem-Solving Approach?

To investigate if the task design motivated students to reflect on their problem-solving approach, it was necessary to check students' awareness of their approach. Accordingly, we asked the participants to describe how they solved the tasks and whether they had reflected about chemical concepts. Haley was one of the few participants aware of her rote learning approach; half of the students answered that they did not think about chemical concepts when solving subtasks A1 and A2.

> *Haley: "It was more like an automatism that I used. Somehow, I saw the reaction and reagents before. That's how I do it. I don't see it and think about how they behave and so on. It's more like "shoot from the hip" as the phrase goes."*

Haley described that solving subtasks A1 and A2 initiated an automation process. She stated that she had recognized explicit surface features of the given structures and recalled the matching product. While processing the task, Haley used a problem-solving approach that revealed her rote memorization approach. Subtasks A1 and A2 contained not only familiar reactions for the students but also intentionally easy ones. Haley's comments show that the task-design leads, as intended, to an activation of rote learning approaches because it is the fastest way to propose a product for the reaction. When students memorize the products of a reaction they tend to ignore byproducts that could be generated as well.[33] Memorizing reaction products often leads to the neglect of

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

underlying chemical processes. Hence, students tend to be easily satisfied with their memorized products instead of being critical of their decision. Subtask A3 is about weighing chemical concepts and estimating the plausibility of alternative pathways. Rote learning approaches may or may not be productive when solving this subtask. After solving A3, Mitchell was asked the same question as Haley; his statement highlights that he valued the more challenging task.

Mitchell: "I think it is favorable to reflect on all the alternatives and therefore even more about the chemical processes that run in the background."

Mitchell realized that he had to reflect on the product cards in A3 using more meaningful techniques because he referred to chemical concepts and mentioned that these processes are not visible at first sight. Mitchell's meaningful learning approach does not guarantee a correct solution, per se, but he seemed to be aware that the subtask required more analytical thinking to lead him to a more appropriate solution. All 29 students confirmed that they had reflected more than usual about chemical concepts while solving A3/B3. Hence, it is desirable that students use meaningful techniques more often, not only to practice weighing chemical concepts but also to become aware that organic chemistry is a process-oriented field.[11] To become aware of their own approach, students need to learn first-hand, that rote learning does not necessarily guarantee success. In this regard, Phil serves as an example of a student who changed his approach because he began to doubt his previous decisions during the task exercise and clearly expressed his cognitive dissonance.

Phil: "I recognized that when you look at the first two subtasks, that is kind of automized. You do the S_N, that's it, and you go on. As I saw the other products, I began to have doubts that made me think twice about it and even go deeper. What alternatives do I have and how plausible are they?"

As was our intention for subtasks A1 and A2, Phil initially used a problem-solving approach that was connected to rote learning. He stated that he solved A1 and A2 rather automatically and did not reflect critically on his answers. In A3, the variety of products initiated his awareness that thinking about plausibility and alternatives might be necessary. By using this rather meaningful problem-solving approach, Phil further debated several chemical concepts and tried to weigh them while solving the subtasks. Like Phil, more than three-quarter of the students solved subtasks A1 and A2 in a rather automatic manner.

Many students raised doubts about their initial problem-solving approach, especially after being confronted with the product cards in A3. Interestingly, not all students valued such an intentionally "enforced" problem-solving approach as positively as Phil. Gloria, one of four such students, claimed that the subtasks evoked uncertainty.

Gloria: "These tasks made me uncertain. I questioned my decisions and I'm a person who tends to lose track of the big picture."

Gloria's uncertainty is understandable, in that students often feel uncomfortable when they do not know the correct solution or when they question their answers, especially in an environment where teachers value the direct recall of information. On the other hand, it is possible that the tasks caused a cognitive overload and therefore created an exhausting situation for students.[34] Gloria might need more practice and further scaffolding to approach the task meaningfully. Even when students express uncertainty, the tasks can be used as a diagnostic tool to identify students who have reached the limit of their abilities with regard to conceptual understanding or problem-solving. Even when students are not certain about their conceptual understanding or do not have the ability to solve the tasks successfully, these tasks can help to focus students on reworking chemical concepts or rethinking their problem-solving approach.

Which Supportive Conditions Do Students Consider Important for Adopting a More Meaningful Approach to Solving Tasks?

To determine how instructions can be designed to support students to reflect more deeply on alternative mechanistic pathways, we asked participants to reflect on supportive conditions. Most students gave similar answers. Cameron's answer is an example of the most prevalent opinion.

Cameron: "The problem is for these kinds of tasks that you need someone to discuss it with. [···] A fellow student would be sufficient for this. It would be enough if he said 'How do you come to this conclusion?' You don't need someone who knows the correct answer, only someone asking 'Why this answer?' I prefer this [approach to problem-solving]."

In general, students said that a partner would be crucial for a valuable discussion, since they were less likely to think about a problem in depth on their own than in discussion with a fellow student. Working with a partner, or cooperative learning, is an essential part of active learning.[35,36] When solving tasks on their own, with a partner or in groups, discussions are more likely to motivate students to engage in counter-arguments and to think beyond the surface to reason out the "chemical processes that run in the background," as Mitchell remarked.[37]

Participants were also asked to compare the tasks presented in the study with the traditional learning environment—a class lecture or exercises in which sample solutions are presented—that they had experienced during their study program. For instance, Claire was asked this question after tasks A and B. In her answer, she compared the tasks to the traditional exercises she had experienced as follows.

Claire: "I haven't done something like this in the last two years because you've just got something in front of you telling you that this is the mechanism of the reaction and that's it. [···] But I should have created a concept that I could refer back to."

Claire emphasized that the traditional environment provided her with the correct reaction mechanisms and did not prompt her to reason out the underlying chemical concepts. About three-quarters of the participants gave similar answers stating that they did not think about the concepts they should have learned in the traditional lecture setting. Claire emphasized her experience with the traditional setting again after task B as follows.

Claire: "During [my studies in] the last few years, things have just been written in the board and you copied it down. You could've just copied the textbook instead."

Claire mentioned that in the traditional setting she only copied sample solutions that were written on the board. In contrast, studies have revealed that students work more intensively and learn more in cooperative learning environments than in traditional setting because active learning approaches provide more opportunities for discussion.[37,38] This could be because traditional settings convey a vast amount of information in a brief time span so that students use

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

the rote learning approach to cope with the deluge of information. In this study, the presented tasks created learning settings in which students had the opportunity for discussion with one another, an activity that can activate analytical thinking. Analytical thinking and the use of chemical concepts are central skills as weighing reactivities and judging reaction centers is, for instance, essential to be aware of alternative byproducts.[7,33] Therefore, tasks used in organic chemistry lectures should focus on supporting students to actively use and develop their understanding of chemical concepts, instead of relying on rote learning approaches. This finding has urgent implications for changes in organic chemistry education,[6,8,19] so that Claire and her classmates can learn to "create a concept that they can refer back to" and apply to other situations. However, creating opportunities for analytical thinking is only the beginning of valuing the development of these skills.

■ CONCLUSION AND IMPLICATIONS

The presented task sequence, focusing on creating cognitive dissonance, was designed to challenge the students' problem-solving approach. It was apparent that students used a rote learning approach with simple, familiar tasks at the beginning of the sequence.[39] When the tasks became more complex, simple recall was not productive in solving these tasks successfully. Thus, students experienced a cognitive conflict when they had to evaluate the plausibility of five alternative products that eventually contradicted their own solutions.[27] This cognitive conflict appeared to have initiated a change in students' problem-solving approach from rote memorization to a more meaningful, analytical approach.[39] Students had to think deeply about the given alternatives to reach a decision.[19] Moreover, due to the given reaction alternatives, the task sequence seemed to have prevented students from focusing on one main product. Instead, the task sequence led student to evaluate the plausibility of byproducts and alternative pathways, which they did by weighing chemical concepts. This is a very important process for planning syntheses and reasoning mechanistically.[33,40] As a result, the tasks created in students an increased awareness of their problem-solving strategies so that they began to question them. Nevertheless, the tasks do require that students have a certain amount of prior knowledge and are able to activate, integrate, and use their knowledge.[41]

The aim of this activity was not to prevent students from rote learning. Rather, its goal was to help students become aware of their problem-solving approach and weigh the pros and cons of rote learning and meaningful learning.[9] With an increased awareness of their problem-solving approach, students may be able to use and apply conceptual knowledge broadly; this can be helpful in the laboratory where rote learning of reaction mechanisms may be ineffective because reagents or conditions can differ and have an effect on the outcome of a reaction. Therefore, these tasks and situations were generated to develop students' ability to apply chemical concepts.

As these tasks can be quite complex, depending on the student cohort, there are different ways to design and use them. When teachers want to design these tasks on their own, they could use their own students' erroneous solutions and diagnose them beforehand. The task could integrate these alternative ideas and conceptions of their respective group of students.

Moreover, designing tasks can also be a student activity in which students can reflect on their own problem-solving approach and also take an outsiders' perspective to create different product cards.[32] On the basis of our analysis, when implementing the tasks in the classroom, students should be encouraged to discuss their answers in groups or with a partner.

In the classroom, active pedagogies such as Think−Pair−Share (TPS) or cooperative methods could be used to facilitate the implementation of the tasks. The tasks can also be used as a diagnostic tool, as an exercise to review chemical concepts (as we did) or as part of an oral exam in which several chemical concepts are integrated.

However, creating opportunities for analytical thinking and reflecting on one's own erroneous approaches to problem-solving is only a first step toward more meaningful learning. If students' meaningful approaches to problem-solving (even if they are erroneous) are not valued in assessments, rote memorization will remain the dominant problem-solving approach taken by students.

■ LIMITATIONS

The tasks in this study were well received by the students and their feedback was consistently positive. However, four students said that the tasks were overwhelming, or that they lacked knowledge of several chemical concepts and so were not able to complete the tasks meaningfully. These students require additional scaffolding on how to approach these complex tasks. They tend to reach a decision very quickly and neglect key variables, which can lead to erroneous solutions. The tasks, particularly with an additional scaffold, can help slow down the decision-making process in, for example, collecting and weighing arguments before reaching a decision.[42] This progressive approach can lead to a recognition of more implicit features to incorporate a greater number of variables.[43]

The impact of these tasks was evaluated using interviews and the data were analyzed both quantitively and qualitatively. However, our conclusions should be considered with caution when implementing the tasks in the actual classrooms. This is because the interviews may have influenced students in the study to use their reasoning and analytical skills more intensively than they would have used it normally.

The tasks were primarily designed as a diagnostic tool to analyze organic chemistry students' problem-solving approach and their difficulties when reflecting on alternative reaction products. However, to what extent the task-design will influence students' reasoning behavior cannot be determined based on our current study.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.0c00248.

Task A/B, corresponding product cards, and sample solutions (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Nicole Graulich − *Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen 35392, Germany;* orcid.org/0000-0002-0444-8609;
Email: Nicole.Graulich@dc.jlug.de

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

Author

Leonie Lieber − *Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen 35392, Germany;* orcid.org/0000-0001-9955-6655

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jchemed.0c00248

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This publication is part of the first author's doctoral (Dr. rer. nat.) thesis at the Faculty of Biology and Chemistry, Justus-Liebig-University Giessen, Germany. We thank the students who participated in the study, the members of the Graulich research group, Krenare Ibraj, Dr. Ira Caspari, and Prof. Dr. Richard Göttlich for their support and fruitful discussions. Leonie Lieber would like to thank the Verband der Chemischen Industrie (German Chemical Industry Association) for supporting her with the Kekulé Fellowship.

■ REFERENCES

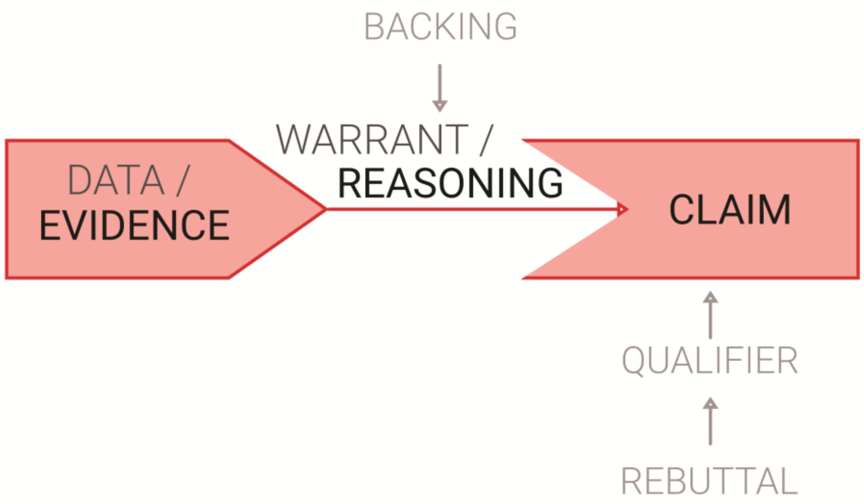
(1) Krathwohl, D. R. A revision of Bloom's taxonomy: An overview. *Theory Pract.* **2002**, *41* (4), 212−218.
(2) Bloom, B.; Engelhart, M.; Furst, E.; Hill, W.; Krathwohl, D. *Taxonomy of educational objectives, the classification of educational goals, Handbook 1: Cognitive domain*; David McKay Company. Inc: New York: 1956.
(3) Flynn, A. B. Developing problem-solving skills through retrosynthetic analysis and clickers in organic chemistry. *J. Chem. Educ.* **2011**, *88* (11), 1496−1500.
(4) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable reasoning: multi-variate problem-solving in organic chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281−292.
(5) Cartrette, D. P.; Bodner, G. M. Non-mathematical problem solving in organic chemistry. *J. Res. Sci. Teach.* **2010**, *47* (6), 643−660.
(6) Bode, N. E.; Deng, J. M.; Flynn, A. B. Getting Past the Rules and to the WHY: Causal Mechanistic Arguments When Judging the Plausibility of Organic Reaction Mechanisms. *J. Chem. Educ.* **2019**, *96*, 1068−1082.
(7) DeCocq, V.; Bhattacharyya, G. TMI (Too much information)! Effects of given information on organic chemistry students' approaches to solving mechanism tasks. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 213−228.
(8) Bhattacharyya, G.; Bodner, G. M. "It gets me to the product": How students propose organic mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402.
(9) Grove, N. P.; Bretz, S. L. A continuum of learning: from rote memorization to meaningful learning in organic chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 201−208.
(10) Talanquer, V., How do students reason about chemical substances and reactions? In *Concepts of matter in science education*; Springer: 2013; pp 331−346.
(11) Grove, N. P.; Cooper, M. M.; Cox, E. L. Does mechanistic thinking improve student success in organic chemistry? *J. Chem. Educ.* **2012**, *89* (7), 850−853.
(12) Evans, J. S. B. In two minds: dual-process accounts of reasoning. *Trends Cognit. Sci.* **2003**, *7* (10), 454−459.
(13) Graulich, N. Intuitive judgments govern students' answering patterns in multiple-choice exercises in organic chemistry. *J. Chem. Educ.* **2015**, *92* (2), 205−211.
(14) Gigerenzer, G.; Gaissmaier, W. Heuristic decision making. *Annu. Rev. Psychol.* **2011**, *62*, 451−482.
(15) Talanquer, V. Commonsense chemistry: A model for understanding students' alternative conceptions. *J. Chem. Educ.* **2006**, *83* (5), 811−816.
(16) Anzovino, M. E.; Bretz, S. L. Organic chemistry students' ideas about nucleophiles and electrophiles: the role of charges and mechanisms. *Chem. Educ. Res. Pract.* **2015**, *16* (4), 797−810.
(17) Anzovino, M. E.; Bretz, S. L. Organic chemistry students' fragmented ideas about the structure and function of nucleophiles and electrophiles: a concept map analysis. *Chem. Educ. Res. Pract.* **2016**, *17* (4), 1019−1029.
(18) Cooper, M. M.; Corley, L. M.; Underwood, S. M. An investigation of college chemistry students' understanding of structure−property relationships. *J. Res. Sci. Teach.* **2013**, *50* (6), 699−721.
(19) Talanquer, V. Concept Inventories: Predicting the Wrong Answer May Boost Performance. *J. Chem. Educ.* **2017**, *94* (12), 1805−1810.
(20) Morewedge, C. K.; Kahneman, D. Associative processes in intuitive judgment. *Trends Cognit. Sci.* **2010**, *14* (10), 435−440.
(21) Graulich, N. The tip of the iceberg in organic chemistry classes: how do students deal with the invisible? *Chem. Educ. Res. Pract.* **2015**, *16* (1), 9−21.
(22) McClary, L.; Talanquer, V. College chemistry students' mental models of acids and acid strength. *J. Res. Sci. Teach.* **2011**, *48* (4), 396−413.
(23) Newell, B. R.; Shanks, D. R. Take the best or look at the rest? Factors influencing" one-reason" decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* **2003**, *29* (1), 53−65.
(24) Glöckner, A.; Betsch, T. Multiple-reason decision making based on automatic processing. *J. Exp. Psychol. Learn. Mem. Cogn.* **2008**, *34* (5), 1055−1075.
(25) Festinger, L. *A Theory of Cognitive Dissonance*. Stanford University Press: 1957; Vol. 2.
(26) Linenberger, K. J.; Bretz, S. L. Generating cognitive dissonance in student interviews through multiple representations. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 172−178.
(27) Afmeur, E.; Dufort, H.; Leibu, D.; Frasson, C. In Some justifications for the learning by disturbing strategy. Artificial Intelligence in Education, 1997: Knowledge and Media in Learning Systems: Proceedings of AI-ED 97, World Conference on Artificial Intelligence in Education, IOS Press, Kobe, Japan: 1997; p 119.
(28) Heemsoth, T.; Heinze, A. The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instr. Sci.* **2014**, *42* (4), 639−657.
(29) Loibl, K.; Leuders, T. How to make failure productive: Fostering learning from errors through elaboration prompts. *Learn. Instr.* **2019**, *62*, 1−10.
(30) Akkuzu, N.; Uyulgan, M. A. An epistemological inquiry into organic chemistry education: exploration of undergraduate students' conceptual understanding of functional groups. *Chem. Educ. Res. Pract.* **2016**, *17* (1), 36−57.
(31) Cartrette, D. P.; Mayo, P. M. Students' understanding of acids/bases in organic chemistry contexts. *Chem. Educ. Res. Pract.* **2011**, *12* (1), 29−39.
(32) Milkman, K. L.; Chugh, D.; Bazerman, M. H. How can decision making be improved? *Perspect. Psychol. Sci.* **2009**, *4* (4), 379−383.
(33) Popova, M.; Bretz, S. L. It's Only the Major Product That We Care About in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams. *J. Chem. Educ.* **2018**, *95* (7), 1086−1093.
(34) Sweller, J. Evolution of human cognitive architecture. *Psychology of learning and motivation* **2003**, *43*, 216−266.
(35) Adams, D.; Carlson, H.; Hamm, M. *Cooperative learning & educational media: Collaborating with technology and each other*; Educational Technology: 1990.
(36) Johnson, D. W.; Johnson, R. T.; Smith, K. The state of cooperative learning in postsecondary and professional settings. *Educ. Psychol. Rev.* **2007**, *19* (1), 15−29.
(37) Tsay, M.; Brady, M. A case study of cooperative learning and communication pedagogy: Does working in teams make a difference? *J. Scholars. Teach. Learn.* **2010**, 78−89.

(38) Carlsmith, K. M.; Cooper, J. A persuasive example of collaborative learning. *Teach. Psychol.* **2002**, *29* (2), 132−135.
(39) Bhatia, S. Conflict and bias in heuristic judgment. *J. Exp. Psychol. Learn. Mem. Cogn.* **2017**, *43* (2), 319.
(40) Caspari, I.; Weinrich, M.; Sevian, H.; Graulich, N. This mechanistic step is "productive": organic chemistry students' backward-oriented reasoning. *Chem. Educ. Res. Pract.* **2018**, *19* (1), 42−59.
(41) Ausubel, D. P.; Novak, J. D.; Hanesian, H., *Educational Psychology: A Cognitive View*; Holt, Rinehart and Winston: 1968.
(42) Graulich, N.; Caspari, I. Designing a scaffold for mechanistic reasoning in organic chemistry. *Chem. Teach. Int. Best Pract. Chem. Educ.* **2020**, DOI: 10.1515/cti-2020-0001.
(43) Caspari, I.; Graulich, N. Scaffolding the structure of organic chemistry students' multivariate comparative mechanistic reasoning. *Int. J. Phys. Chem. Educ.* **2019**, *11* (2), 31−43.

https://dx.doi.org/10.1021/acs.jchemed.0c00248
J. Chem. Educ. 2020, 97, 3731−3738

3 Investigating Students' Argumentation when Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry

In Chapter 3, the original paper is represented. It is reproduced from Lieber and Graulich (2022) with permission of the Royal Society of Chemistry.



Lieber, L. & Graulich, N. (2022). Investigating Students' Argumentation when Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry. *Chemistry Education Research and Practice*, 23, 38-54.

Chemistry Education Research and Practice

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Chem. Educ. Res. Pract.*, 2022, **23**, 38

Investigating students' argumentation when judging the plausibility of alternative reaction pathways in organic chemistry

Leonie Lieber and Nicole Graulich *

Received 28th May 2021,
Accepted 23rd August 2021

DOI: 10.1039/d1rp00145k

rsc.li/cerp

Building scientific arguments is a central ability for all scientists regardless of their specific domain. In organic chemistry, building arguments is a necessary skill to estimate reaction processes in consideration of the reactivities of reaction centres or the chemical and physical properties. Moreover, building arguments for multiple reaction pathways might help students overcome the tendency toward one-reason decision-making and offer them an authentic perspective on organic processes. Reasoning about multiple alternative organic reaction pathways requires students to build arguments and then judge and weigh the plausibility of these pathways. However, students often struggle to build strong arguments and use scientific principles appropriately to justify their claims. In the present study, the argumentation patterns of 29 chemistry majors students were analysed using a simplified version of Toulmin's argumentation model (claim–evidence–reasoning). The students solved various tasks related to alternative reaction pathways of a substitution reaction. They supported their claims with evidence and justified the evidence through reasoning. We investigated (a) the extent to which the students use evidence and reasoning in their argumentation (referred to as their argumentation approach), (b) how students with different argumentation approaches rationalised changes in their initial claims, and (c) how students used reasoning to justify their arguments. The results indicate that students need further support to appropriately use evidence and reasoning and to apply conceptual knowledge to build well-grounded arguments.

Introduction

In scientific research, sequencing a complex problem into manageable pieces simplifies the problem, and this approach is often referenced in the scientific method described by Francis Bacon (Bacon, 1878; Franck, 2012). The scientific method involves the formulation of a hypothesis, followed by testing it to verify or falsify the hypothesis (Turro, 1986). Almost 60 years ago, Platt (1964) emphasised the importance of considering alternatives as part of the scientific method. In Platt's perspective on the scientific method, which is called strong inference, the aforementioned steps of the scientific method are enhanced through the creation of alternative hypotheses. Platt (1964) claimed that the explicit and regular use of alternatives at each step of the research process offers additional power. Turro (1986) reported that considering alternatives allows a researcher to stay open-minded to deviations or unexpected findings rather than focused exclusively on one hypothesis and possibly missing interesting findings. However, incorporating alternatives into decision-making can be challenging. One way to support this decision-making process is to build well-grounded arguments for various alternatives.

Justus-Liebig University Giessen, Institute of Chemistry Education, Heinrich-Buff-Ring 17, 35392 Giessen, Germany. E-mail: nicole.graulich@dc.jlug.de

Building arguments or explanations is a core ability that must be learned by science students. This process includes evaluating claims and weighing evidence (Driver *et al.*, 2000). Erduran (2019) purposefully linked these two aspects (claims and evidence), which were previously considered separate, by stating that claims need to be justified with evidence. Justifying and refuting claims using evidence is part of a scientific mode of thinking.

In organic chemistry, a student's success is often linked to the ability to build arguments (Bodé *et al.*, 2019; Cruz-Ramirez de Arellano and Towns, 2014; Deng and Flynn, 2021) and reason about reaction mechanisms (Grove *et al.*, 2012; Becker *et al.*, 2016; Caspari *et al.*, 2018; Watts *et al.*, 2021). More explicitly, this means that students who solve tasks related to reaction mechanisms need to be able to argue and reason about the structural changes that occur and the causes of these changes. To explain these causes, students need to justify their evidence with reasoning (Cooper *et al.*, 2016; Watts *et al.*, 2020). However, students show various problems when building arguments and using reasoning. Students often struggle to consider alternative pathways since they tend to focus on one reaction product instead of the underlying processes (Popova and Bretz, 2018a). Thus, students are often triggered by single clues such as the presence of a leaving group, which often leads them to neglect other reactivities or multiple reaction pathways (Kraft *et al.*, 2010). With regard to argumentation structures, forming a claim is often the easiest component for students (McNeill and Krajcik, 2012). Problems occur when students try to justify claims as they experience uncertainties regarding what counts as evidence (Sadler, 2004). Thereby, students explicitly struggle with the appropriate use of scientific principles while explaining why their given evidence supports a claim (McNeill *et al.*, 2006; McNeill and Krajcik, 2012; Walker *et al.*, 2019). Problems in students' understanding of chemical concepts such as nucleophilicity and electrophilicity or acidity and basicity are well documented (Cartrette and Mayo, 2011; Anzovino and Bretz, 2015; DeFever *et al.*, 2015; Akkuzu and Uyulgan, 2016). To avoid these problems in chemical concepts, students rely on personal views when judging the plausibility of a given context (Hogan and Maglienti, 2001). Even when students justify their claims, they often rely on one-reason decision-making based on single pieces of evidence when multiple supporting pieces are needed (McNeill *et al.*, 2006; Talanquer, 2006; Kraft *et al.*, 2010). This is primarily observed when students must argue about data that contradict their own ideas (Chinn and Brewer, 2001).

Use of argumentation models in chemistry education research

Research in the field of chemistry education has provided information on how to support students' argumentation skills through assessments, activities, or teaching approaches (Cooper, 2015; Talanquer, 2018; Bodé *et al.*, 2019; Luo *et al.*, 2020). Studies on the formation of arguments by students have been conducted in the fields of natural science (Abi-El-Mona and Abd-El-Khalick, 2011; Luo *et al.*, 2020), laboratory work (Hand and Choi, 2010; Walker *et al.*, 2019; Hosbein *et al.*, 2021; Petritis *et al.*, 2021), physical chemistry (Becker *et al.*, 2013; Towns *et al.*, 2019), and organic chemistry (Cruz-Ramirez de Arellano and Towns, 2014; Bodé *et al.*, 2019; Deng and Flynn, 2021). Some studies purposefully used contrasting cases to capture students' argumentation abilities in terms of weighing properties, reaction centres, or reaction conditions (Caspari *et al.*, 2018; Bodé *et al.*, 2019; Galloway *et al.*, 2019; Graulich *et al.*, 2019; Watts *et al.*, 2021).

Many chemistry education researchers have specifically used Toulmin's argumentation model to analyse or support argumentation. For instance, Becker *et al.* (2013) analysed classroom discourse in which students made claims and particulate-level justifications about thermodynamic topics such as enthalpy and heat capacity. The authors suggested that while a conceptual understanding of the particular nature of matter is important, students must to be able to apply this knowledge about chemical and physical properties to build arguments (Becker *et al.*, 2013). Lazarou and Erduran (2021) analysed science teachers' instructional adaptions of Toulmin's argumentation pattern in the classroom and the quality of students' written arguments. They revealed that some elements of Toulmin's argumentation pattern cause ambiguity and interpretative difficulties for students and science teachers (Lazarou and Erduran, 2021).

The claim–evidence–reasoning (CER) model, a simplified version of Toulmin's argumentation model, is currently the prevalent argumentation model used in chemistry education research. Walker *et al.* (2019) conducted a study in a general chemistry laboratory course based on the argument-driven inquiry instructional model. The participants worked in groups and collected data in the laboratory to build arguments that included claims, evidence, and reasoning as justifications. The authors found that students had problems changing their claims and considering alternatives. Instead, they stuck with their initial claims despite critique. Moreover, students struggle to justify their claims with scientific principles. Walker *et al.* (2019) reported that students may need additional guidance to revise or change their claims and to use scientific principles in their justification. Luo *et al.* (2020) investigated the effect of inquiry on students' argumentation skills. They examined the influence of a reasoning flow scaffold in comparison to a control group that received conventional argumentation training. The students in the intervention group showed significant improvement in certain argumentation elements compared to the control group. Luo *et al.* (2020) concluded that dividing a claim into two parts (one at the beginning and one at the end as an opportunity for falsification) may assist students in building arguments by causing them to reflect on their decision more carefully. In organic chemistry, Bodé *et al.* (2019) investigated how students build scientific arguments when comparing two reaction mechanisms for nucleophilic substitution. They concluded that practicing argumentation skills are essential to make expectations clear for the students. These practices may include debating about justifications or comparing with prior knowledge.

Overall, the above studies indicate that students (1) struggle to revise their initial, possibly erroneous claims, (2) rarely consider alternatives in their argumentation, and (3) struggle to build well-grounded arguments.

Clarifying how students use evidence and reasoning in their arguments and rationalise their claims when assessing the plausibility of multiple alternative reaction pathways may provide further insights into their argumentation behaviours.

Theoretical framework

Toulmin's argumentation pattern

Toulmin can be considered as one of the "founding fathers of modern argumentation theory" because of the impact his ideas have on everyday reasoning (Van Eemeren *et al.*, 2014: p. 204).

Toulmin determined that an argument consists of six elements, as represented in Fig. 1, which also illustrates the commonalities between Toulmin's argumentation pattern and the simplified CER model (Toulmin, 2003). More generally, Kuhn (1991) claimed that an argument is an assertion with justification, while Osborne and Patterson (2011) stated that an argument is a justified claim.

The following descriptions of the elements of an argument are taken from Toulmin's book *The Uses of Argument* (Toulmin, 2003).

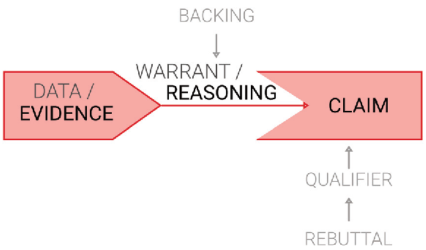


Fig. 1 Coherence of Toulmin's argumentation pattern (2003) and the CER model often used in chemistry education research.

The first step in building an argument is the expression of a claim. Asserting something is the key component of making a claim (Van Eemeren *et al.*, 2014). The foundation of the claim is called data. Data are initiated by the question "What have you got to go on?". Therefore, the claim based on the given data is the first step to justify the claim (and the second step of an argument). The given data consist of rules or principles and show that the claim is appropriate. The third step of an argument is called the warrant, which is the justification of the given data. The warrant is a general statement based on rules or principles. Since a warrant is a bridge between the data and the claim, the initiating question is "How do you get from your data to your claim?". On the basis of data, the warrant gives permission to make a claim. The fourth element of an argument is the backing, which is based on the question "What entitles you to conclude from... to...?" and refers to the warrant. A backing is required when the warrant is not immediately accepted. Thus, the backing relies on the principles of ethics, values, or general norms. The fifth element of an argument is called the qualifier and is initiated by the question "Is that necessarily so?". The qualifier provides information about the obligation of the claim and thus possibly weakens the claim. Words such as likely, probably, or necessarily can be applied to introduce a qualifier. The sixth and last element of an argument is the rebuttal. The rebuttal describes the circumstances under which the claim is not valid or only valid to a certain extent. A rebuttal is an exception of rules and is initiated with the question "When does the rule not apply?". As soon as a rebuttal appears in an argument, the claim must be weakened with a qualifier. Conversely, just because there is a qualifier, there is not necessarily the need for a rebuttal (Toulmin, 2003).

Claim–evidence–reasoning (CER) model

Toulmin noted that of the six aforementioned elements, only the claim, data, and warrant are present in every argument, Thus, these three elements form the core of the argument. The rebuttal, qualifier, and backing are only present when they are needed (Toulmin, 2003).

Accordingly, CER is frequently used as a simplified version of Toulmin's argumentation pattern based on the terms "claim," "evidence" (data), and "reasoning" (warrant).

In both the original and CER models, the main goal of an argument is the justification of a claim with the use of evidence and reasoning (Toulmin, 2003; Osborne and Patterson, 2011). The CER model consists of three parts (McNeill and Krajcik, 2012).The first part in both the CER and original models is the claim, which acts as a statement or a conclusion to a problem or question (McNeill *et al.*, 2006; McNeill and Krajcik, 2012). Similar to Toulmin's argumentation pattern, the claim in the CER model is always in doubt and requires a justification based on known data and an explanation (Osborne and Patterson, 2011). The claim is supported with evidence based on scientific data (McNeill *et al.*, 2006; McNeill and Krajcik, 2012). Thus, evidence in den CER model is similar to Toulmin's data (Van Eemeren *et al.*, 2014). As scientific data contain information from various resources, multiple pieces of evidence can support a single claim (McNeill and Krajcik, 2012). Equally important is the reasoning (in the CER model) as it provides the bridge between the claim and the evidence; reason in the CER model combines Toulmin's warrant and backing (McNeill *et al.*, 2006; McNeill and Krajcik, 2012). Reasoning indicates how the link between a claim and the given data can be justified and why they are justified by providing a logical connection (warrant) (McNeill and Krajcik, 2012; Toulmin, 2003; Van Eemeren *et al.*, 2014). As this connection may need backing up, scientific principles should be included (backing) (McNeill and Krajcik, 2012). Reasoning is considered to be the most difficult step of an argument because a justification requires the bridging of the claim and evidence (McNeill and Krajcik, 2012).

Research questions

Given the need to diagnose how students build arguments in organic chemistry, our goal was to investigate students' use of evidence and reasoning (*i.e.*, the CER argumentation pattern) when forming arguments about the plausibility of alternative reaction pathways. We separated our overarching question into three research questions:

(1) To what extent do students use evidence and reasoning in their argumentation (which we refer to as the argumentation approach)?

(2) How do students with different argumentation approaches rationalise changes in their initial claims?

(3) How do students use reasoning to justify their arguments while assessing the plausibility of alternative reaction products?

To answer these three research questions, we conducted qualitative interviews with students majoring in chemistry prompting them to elaborate on serial tasks related to the plausibility of reaction products.

Methods

Context and study setting

The study was conducted at a German university in October and November of 2019. Twenty-nine participants were recruited on a voluntary basis at the beginning of the course "Organic Chemistry 3 – Catalysis and Synthesis" (OC3), which is part of the bachelor of science chemistry major programme in the fifth of six semesters. Before students can attend Organic Chemistry 3, they must pass the Organic Chemistry 1 (OC1) and Organic Chemistry 2 (OC2) courses. Both courses consist of a lecture, a seminar, and a laboratory section. These two organic chemistry courses cover topics such as the reactivities of functional groups, structure–property relationships, organic reaction mechanisms (*e.g.*, nucleophilic substitutions, rearrangements, and carbonyl reactions), hard and soft acid and base (HSAB) theory, and transition state theory. To successfully complete these two courses, students must pass a written exam. The aforementioned course topics covered in OC1 and OC2 were considered to be prior knowledge for the OC3 students as the study was conducted during the first four weeks of the OC3 course.

Eleven female and 18 male chemistry major students participated on a voluntary basis. Students were recruited at the beginning of the OC3 course *via* an announcement in the lecture. The students were between 20 and 27 years old (average 22 years). While Institutional Review Board is not required at German universities, interviews followed ethical guidelines, and the students had the opportunity to terminate their interview at any point. All recruited students were provided with information about their rights and the manner of data collection and gave their written permission for the following: (1) the questionnaire can be evaluated by the research team, (2) the transcripts, drawings, video data, and audio data can be used by the research team; and (3) the collected data can be analysed and published by the research team. All students created a pseudonym, and the video data only included voice and drawings. As all students were native German speakers, the interviews were conducted in German, and direct quotes were translated into English for publication.

Data collection

During the interviews, the students were encouraged to draw products and reaction mechanisms to support their statements. Most of the students used drawings to solve the tasks.

The students completed a demographic questionnaire at the beginning of the interview. The data also included pseudonymised scans of the students' work sheets as well as audio and video recordings. Verbal utterances were transcribed verbatim. Individual think-aloud interviews were led by the first author using a semi-structured protocol (Bernard and Bernard, 2013). After their interviews, all students were asked not to talk about the tasks with their colleagues to minimise the chances of other students preparing beforehand. After the interviews, the first author provided sample solutions of the task during the lecture.

The quotes used in this publication were checked multiple times by the research group and a German English student teacher to guarantee the validity of the translations.

Research instrument

The research instrument used in the interviews consisted of four subtasks (see Fig. 2); a more detailed description of the task design can be found in Lieber and Graulich (2020).

The first subtask contained a typical reaction that the students had experienced several times throughout the organic chemistry courses and laboratories (see Fig. 2, task 1). This reaction was chosen to provoke the use of intuitive heuristics since nucleophilic substitution and elimination were familiar

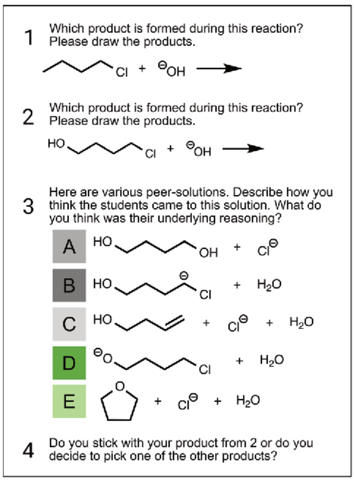


Fig. 2 Illustration of the research instrument.

to the students. The second subtask (Fig. 2, task 2) contained a reaction that – at first sight – seems similar to the reaction from subtask 1 as it only differs in one surface feature. The aim was for the students to use the same problem-solving approach as in subtask 1; however, in this case, the same problem-solving approach would not be productive. In subtask 3 (Fig. 2, task 3), students were provided with alternative products for the reaction given in subtask 2. This step was intended to challenge students' problem-solving approaches as the confrontation with alternative products might initiate a more analytical thinking process. This stepwise process of first generating possibly erroneous solutions by themselves and then comparing their solutions with comparable or different solutions (Loibl and Leuders, 2019) has been shown to have a beneficial effect on students' reasoning. In the last subtask (Fig. 2, task 4), students had the opportunity to defend or revise their generated product from subtask 2 (*i.e.*, revise their initial claim).

Data analysis

All 29 interview transcripts were analysed using the software MAXQDA. All parts of the data analysis were discussed multiple times with the authors and the research group. The data were analysed in three consecutive steps to answer the research questions.

Step I: Identifying students' argumentation patterns

In the first step, we created argumentation pattern maps based on students' interview transcripts to analyse students' argumentation patterns. As students were not familiar with building arguments or the CER pattern, we did not provide them with specific prompts to form evidence and reasoning statements, we rather paraphrased it with why-questions and prompts to explain their reasoning throughout the interview. Prompts were therefore verbalised in a way that students' answers could be assigned to claim, evidence, and reasoning without using the explicit terms. First, we identified the claim as the position being argued for each reaction product card and each student in subtask 4. Two different claims were possible: plausible or implausible. On the basis of the claim, evidence and reasoning were ascertained. To determine what counted as evidence and reasoning, the structure of the statements is important, since a statement can be both evidence and reasoning. Therefore, we prompted the students to provide evidence by having the interviewer ask why the students think something is plausible/implausible. Students responses that were coded as evidence included statements such as "chloride is a good leaving group" or "the H atoms are not acidic". Reasoning was coded when the statement given by the student was a justification of evidence they mentioned before. This was prompted, for example, with the question why a certain leaving group is good or bad. Students answers for this question were, for instance, "chloride is stable" or "C–Cl bond is weaker compared to C–O bond". As students' argumentation patterns differed for each reaction product card, the instances of evidence and reasoning varied for each student and each product card (see Appendix 1).

Step II: building reasoning categories

To obtain insights into the concepts students used while justifying the reaction product cards, we analysed the concepts students referred to in their reasoning. Towards this end, we created eight inductive reasoning categories. Table 1 shows an overview of the final coding rubric for the reasoning codes. We did not consider whether the given statements were technically correct because the aim was to analyse students' argumentation patterns rather than their correctness. Thus, it was possible for technically incorrect statements to be part of the argumentation pattern. A sample solution of all product cards and the corresponding mechanisms has been published by Lieber and Graulich (2020).

A special note is required for the coding of the *Stability* category. This code was only assigned if the students could not explain the term 'stability' after being prompted and relied on judgments such as "the reaction product is stable".

Step III: determining students' argumentation approaches

We then examined the frequency of evidence and reasoning as well as the ratio between them to identify different argumentation approaches. As the claim only consisted of stating if the product card was plausible or implausible, the ratio was calculated separately from the claim. For this reason, claims were omitted from this part of the analysis. To determine the argumentation approaches, both the frequency of statements and the ratio of reasoning to evidence were considered. Therefore, all evidence and reasoning statements across all product cards were aggregated for each student. Based on the results, we grouped the students into three categories according to their argumentation approaches. Group 1 contained students for which the ratio of reasoning to evidence statements was greater than 1 (*i.e.*, the students used more reasoning than evidence in their argumentation pattern). Moreover, the frequency of evidence statements in this group was more than 10, while the frequency of reasoning statements was equal to or greater than 20.

Students were placed into group 2 when the ratio of reasoning to evidence statements was close to 1, and each of the frequencies of evidence and reasoning was between 10 and 20.

Lastly, group 3 contained students for which the ratio of reasoning to evidence was less than or equal to 1 (*i.e.*, the students mostly used evidence than reasoning). For these students, the frequency of evidence statements was nearly 10, while that of reasoning statements was less than 10.

Results and discussion

We analysed students' interview data and their written work to (1), investigate the extent to which students used evidence and reasoning in their argumentation, (2) analyse how students with different argumentation approaches rationalised changes in their claims, and (3) identify how students used reasoning to justify their arguments while assessing the plausibility of reaction products.

To what extent do students use evidence and reasoning in their argumentation?

To answer this question, we analysed the frequencies of evidence and reasoning statements along with their ratio for

Table 1 Illustration of the reasoning categories

Category	Description of the code	Student example
Electronics	Students describe electronic aspects like polarisation, electronegativity, or charge	*"This bond is more polarised."*
Energetics	Students describe energetic aspects like bonds, energies, or thermodynamics	*"The C–H bond is stronger. You have to spend energy to cleave the bond."*
Kinetics	Students describe kinetic aspects like rates or statistics	*". . . that acid–base reactions react very fast."*
Spatial arrangement	Students describe spatial aspects like steric effects, sizes, distances, or angles	*". . . because the groups are close to each other."*
Analogies	Students describe examples that are not part of the task but compare or demonstrate aspects	*"You can deprotonate phenols with hydroxide."*
Strength	Students describe, for instance, the quality or strength of acids/bases, leaving groups, or nucleophiles	*". . . that this part is more acidic."*
Conditions	Students describe aspects like concentrations, temperatures, or pressures	*"It depends on how much base I add."*
Stability	Students describe aspects of stability without an explicit explanation of the term (*e.g.*, in terms of energetics) even after being prompted to do so	*". . . this product would be definitely not stable."*

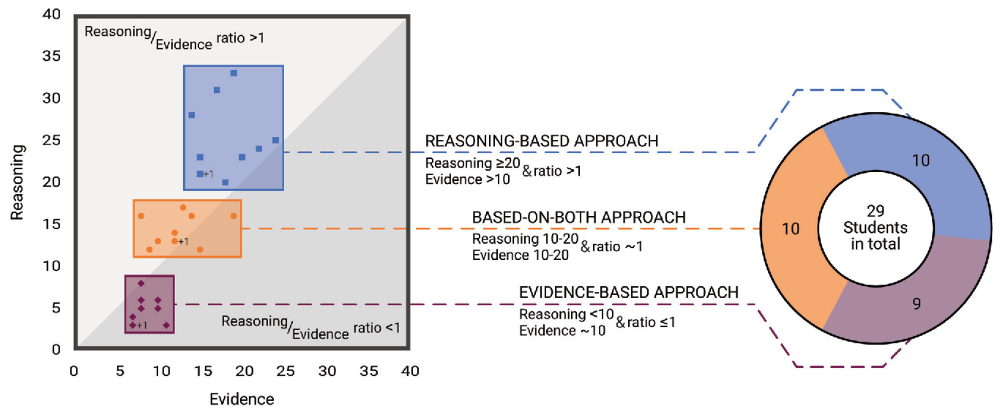


Fig. 3 Classification of the reasoning-based argumentation approach (blue), the based-on-both argumentation approach (orange), and the evidence-based argumentation approach (purple). The diagonal bisector represents a reasoning to evidence ratio of 1. In the left side of the figure "+1" is added to indicate that two students had the same number of evidence and reasoning statements in their argumentation.

each student. We then identified three argumentation approaches that differ based on the evidence and reasoning frequencies and ratio. A detailed description of the students' evidence and reasoning statements can be found in Appendix 1. Fig. 3 illustrates the distribution of students among the three groups of argumentation approaches based on the frequencies of reasoning and evidence. The diagonal bisecting line in Fig. 3 represents a ratio of 1. In addition to the description in the Data analysis section, the classification of students into the three argumentation approaches are detailed as follows.

Argumentation approach 1 (reasoning-based approach): students' argumentation patterns were classified as argumentation approach 1 when the ratio of reasoning to evidence statements was above 1. Additionally, the frequency of evidence statements had to be more than 10, and the frequency of students' reasoning had to be equal to or greater than 20. Both the frequencies of students' evidence and reasoning statements and the ratio has to be accomplished to be assigned to the approach. Ten out of the 29 students were assigned to the reasoning-based argumentation approach.

Argumentation approach 2 (based-on-both approach): students were assigned to argumentation approach 2 when the ratio of reasoning to evidence statements was nearly 1. Moreover, each of the frequencies of evidence and reasoning statements had to be between 10 and 20. The frequencies of evidence and reasoning statements as well as the ratio has to be accomplished to be assigned to the second argumentation approach. Ten out of the 29 students were assigned to this group.

Argumentation approach 3 (evidence-based approach): students were assigned to approach 3 when the frequency of evidence in students' argumentation was nearly 10 while the number of reasoning was lower. The ratio of reasoning to evidence was less than or equal to 1. In contrast to the based-on-both argumentation approach, the frequency of reasoning statements must be lower than 10 and the ratio of reasoning to evidence statements must not be greater than 1 to be assigned to the evidence-based argumentation approach. Thus, nine out of the 29 students were assigned to this approach.

Fig. 4 illustrates example argumentation patterns for the alternative reaction product card C ("alkene") in subtask 3 (see Fig. 2). A complete argumentation pattern map for each argumentation approach is shown in Appendix 2. In each argumentation pattern, the claim (plausible or implausible) and, when present, the support by evidence and the justification by reasoning are shown moving from top to bottom.

All three students whose argumentation patterns are shown in Fig. 4 for product card C initially provided the same correct claim: the alkene is implausible as a reaction product. However, the students differed in the way they supported the claim with evidence or justified it with reasoning. For instance, Reuben, who was assigned to the reasoning-based argumentation approach, supported his claim with three pieces of evidence and justified each piece of evidence with at least one reasoning statement. The reasoning statements were classified into different reasoning categories including *Electronics*, *Spatial Arrangement*, and *Strength*.

Sydney, who was classified into the based-on-both argumentation approach, supported her claim with three pieces of evidence, just as Reuben did. However, unlike Reuben, she did not support each of her evidence statements with reasoning statements.

Frank, who was classified into the evidence-based argumentation approach, provided the same claim as Reuben and Sydney. However, although he supported his claim with a piece of evidence, he was unable to justify it with reasoning even after being asked several times to do so. These three argumentation approaches are illustrated in more detail in the following section with examples provided for each approach.

Fig. 5 illustrates the reasoning-based argumentation approach with a quote from Phil. In general, Phil supported his five claims

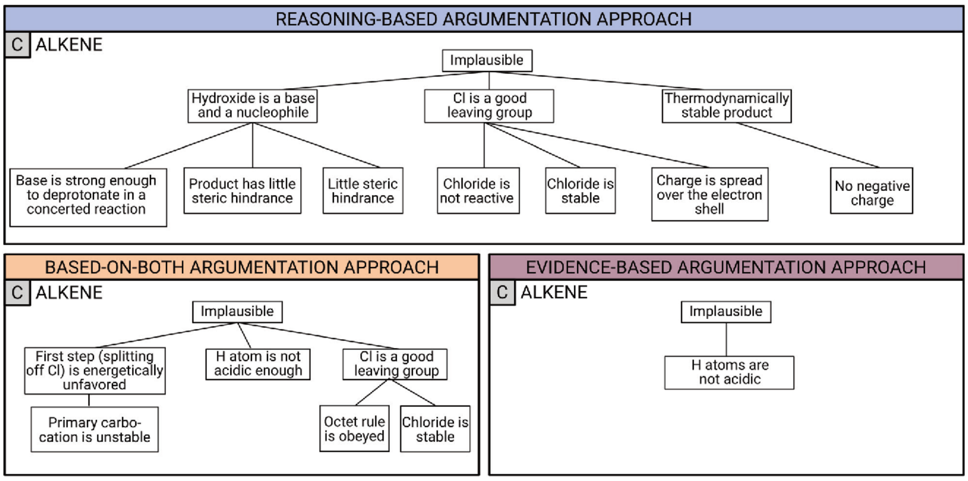


Fig. 4 Example argumentation patterns for the alternative reaction product card C "Alkene": (top) argumentation pattern of student Reuben (reasoning-based argumentation approach); (bottom left) argumentation pattern of student Sydney (based-on-both argumentation approach); and (bottom right) argumentation pattern of student Frank (evidence-based argumentation approach).

for the five reaction product cards with a total of 14 pieces of evidence and 28 reasoning statements (see Appendix 1). In the shown interview excerpt, he argued for reaction product card D, the reaction of 4-chlorobutanol and hydroxide to alkoxide and water, claiming correctly that the reaction product is plausible. He then supported his claim with two pieces of evidence ("If only a very small amount of base is added" and "the acid–base reaction is much faster [compared to the nucleophilic substitution]"). These two statements were categorised as evidence since they refer directly to the claim. Furthermore, he used reasoning to describe statements based on chemical concepts such as electronic or energetic aspects that support the evidence (*e.g.*, why an acid–base reaction is faster compared to a S_N reaction).

To further justify the correct evidence "the acid–base reaction is much faster", Phil provided four reasoning statements. First, Phil used the HSAB principle by reasoning that the hydroxide ion is a hard base, while the proton of the hydroxyl group is a hard acid (*Strength* category). Second, he emphasised that the energy gain is best for the reaction between a hard base and a hard acid (*Energetics* category). Third, Phil referred to the charge density that can be shifted into an empty orbital in the acid–base reaction instead of an antibonding orbital in the nucleophilic substitution (*Electronics* category). Finally, he referenced steric effects related to the ability of the hydroxide ion to attack from any direction in the acid–base reaction (*Spatial Arrangement* category). Overall, Phil used many reasoning statements to justify his evidence, which in turn supported his claim regarding the reaction product card. Thus, Phil was assigned to the reasoning-based argumentation approach.

Dylan serves as an example of the based-on-both argumentation approach. He gave 15 pieces of evidence in total and 12 reasoning statements (see Appendix 1). An interview excerpt of his argument regarding reaction product card E, the reaction of 4-chlorobutanol and hydroxide to tetrahydrofuran (THF) and water, is provided in Fig. 6. Dylan was asked if the reaction product card was plausible or implausible. He mentioned evidence five times correctly ("here [α-carbon atom] is a partial positive charge", "there [hydroxide] is a negative charge", "just a little ring strain", "five-membered ring is quite stable", and "chloride is a good leaving group"). However, only one piece of evidence ("chloride is a good leaving group") was justified with three reasoning statements.

REASONING-BASED ARGUMENTATION APPROACH

Phil: The argument for this would be, **if only a very small amount of base is added, probably this (alkoxide) will be formed**. If we have a 1 to 1 conversion, this would probably come out because **we don't do S_N2 because the acid-base reaction is much faster**. Rather everything is simply deprotonated quantitatively. So I can understand that in any case.

I: You already said the acid-base reaction is way faster than the S_N. Why?

Phil: Yes, that has something to do with hardness. **Here we have a hard base and a proton is a hard acid and according to the HSAB principle the energy gain is best there. I have a positive charge, so I have an empty orbital in which I can push charge density. With a S_N2 I have an antibonding orbital into which I can insert my charge. But there are also steric reasons. It can only attack from there, that alone is a factor, because this is completely terminal, it could take the proton from practically anywhere, and here it must attack specifically.**

Fig. 5 Phil's argumentation for the alkoxide (product card D) as an example for the reasoning-based argumentation approach.

BASED-ON-BOTH ARGUMENTATION APPROACH

I: Do you think this step is plausible?

Dylan: Yes, because **here is a partial positive charge**, **there is a negative charge**. There is also **just a little ring strain**. **A five-membered ring is quite stable** and **chloride is a good leaving group**.

[...]

I: Could you try to explain why chloride is a good leaving group?

Dylan: **Chloride has an electron configuration** and **a high electronegativity**. **Compared to carbon, chloride is more electronegative**. [...] **As an ion, chloride is thermodynamically more stable**.

Fig. 6 Dylan's argumentation for THF (product card E) as an example for the based-on-both argumentation approach.

First, he used reasoning in the *Electronics* category when he said that chloride has an octet. The second reasoning statement is categorised in the *Electronics* category because he indicated that chlorine has a higher electronegativity than carbon. The final reasoning statement was related to the thermodynamic stability of chloride ion (*Energetics* category). Dylan was assigned to the based-on-both argumentation approach since he only supported one piece of evidence with reasoning.

Amber serves as an example of the evidence-based argumentation approach. She used a total of seven pieces of evidence and four reasoning statements to support her claims (see Appendix 1). In the interview, she argued the plausibility of product card A, the reaction of 4-chlorobutanol and hydroxide to 1,4-butanediol and chloride. Here as well, Amber was asked if she thought the reaction product to be plausible or implausible. She answered erroneously that the reaction product is plausible and supported her claim with the evidence that chloride is a good leaving group. Fig. 7 shows an excerpt of Amber's explanation for why she thinks chloride is a good leaving group. At first, she answered that she has no idea why chloride is a good leaving group and that she learned or accepted the fact without questioning it. After the interviewer prompted her to try to explain her evidence, she was quiet for a while before responding incorrectly that chloride could be smaller than hydroxide ion but that she did not know the correct answer. This reasoning statement was assigned to the *Spatial Arrangement* category. Amber was one out of nine students assigned to the evidence-based argumentation approach because she hardly supported her claims with any evidence and barely used reasoning to justify her evidence.

EVIDENCE-BASED ARGUMENTATION APPROACH

I: You have talked about good and bad leaving groups. Why do you think that chloride is a better leaving group than hydroxide?

Amber: I have no idea (laughing).

I: Is this a fact you just know?

Amber: Exactly, I learned that at some point and either didn't question it or simply accpeted it.

I: Could you try to explain it?

Amber: **Maybe because chloride is smaller than hydroxide** but I don't know it.

Fig. 7 Amber's argumentation for the diol (product card A) as an example for the evidence-based argumentation approach.

The three argumentation approaches identified illustrate that the students differed in the amount of evidence and reasoning statements used in their arguments. The more evidence and reasoning a student uses, the higher the probability that a large amount of influential variables are included in the decision-making process. This means that the probability of considering different chemical concepts such as kinetics, energetics, and energetic processes also increases. Since the course of a chemical reaction does not only depend on one factor but rather on various influencing factors, the inclusion of a larger number of evidence statements results in a more comprehensive decision-making process by increasing the breadth of the argumentation process. The number of reasoning statements can then influence the quality or depth of the argument. The more reasoning statements a student uses, the better the evidence can be justified. Reasoning serves to strengthen the claim through the justification of evidence; thus, the students' arguments gain depth and are not just "empty envelopes".

How do students with different argumentation approaches rationalise changes in their initial claims?

We were further interested in how students rationalise a change in their initial claim made in task 1 when prompted to reconsider it in task 4 and if this rationalisation is related to their argumentation approach.

Each student made a claim about each of the five reaction product cards he or she discussed during the interview. Fig. 8 summarises "most plausible" claims made by the students; claims are given as percentages and separated into the three argumentation approaches. In task 2, students formed products independently as they only received the reactants. In contrast, in task 3, five cards with alternative reaction products were provided. All products that were claimed to be "most plausible" by the students in task 2 and task 3 were aggregated so that the total represents 100%. Since the students were able to claim more than one product as plausible, the number of products differs for the tasks and argumentation approaches. For example, 10 students in the reasoning-based argumentation approach claimed the formation of 14 products in task 2 and 15 products in task 3, which corresponded to product cards A-E. The 10 students with the based-on-both argumentation approach built 15 products in task 2 and 21 products in task 3. The nine students with the evidence-based argumentation approach claimed the formation of 12 products in task 2 and nine products in task 3. Regardless of the argumentation approach, most claimed incorrect products such as the diol (product card A) and the alkene (product card C) in task 2. This outcome is not surprising since tasks 1 and 2 were purposely designed to provoke the use of the same problem-solving approach used by

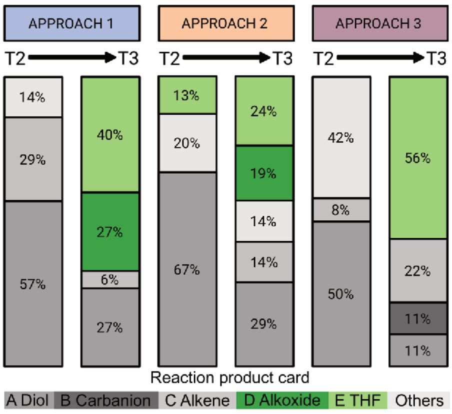


Fig. 8 Total of all "most plausible" claims made in task 2 (T2) and task 3 (T3) separated into the reasoning-based argumentation approach (1), the based-on-both argumentation approach (2), and the evidence-based argumentation approach (3). Students were able to claim more than one product as plausible. "Others" refers to incorrect products such as aldehydes or ketones.

students in task 1, leading to an S_N2 reaction (Lieber and Graulich, 2020). After discussing the reaction product cards in task 3, an increase in claims for the correct product THF (product card E) and its precursor (alkoxide, product card D) along with a decrease in claims of the incorrect product diol (product card A) were observed for all argumentation approaches. Nevertheless, there were noticeable differences in how students with different argumentation approaches rationalised the changes in claims.

Considering only the percentage of THF (product card E) chosen as the most plausible product, it seems that the students who used an evidence-based argumentation approach claimed more correct products (56% THF) compared to those who used a reasoning-based argumentation approach (40% THF) or a based-on-both argumentation approach (24% THF). However, it should be noted that the alkoxide (product card D) is the precursor of THF. Many students who applied reasoning-based and based-on-both argumentation approaches recognised this fact and chose both THF and the alkoxide as the most plausible reaction products. When looking at the total of THF (product card E) and alkoxide (product card D), it is apparent that the students who used an evidence-based argumentation approach did not consider alkoxide as a plausible product or precursor. For example, Charlie who applied a reasoning-based argumentation approach chose THF and the precursor alkoxide as the most plausible products in task 3 when asked if he thought THF is plausible or implausible.

Charlie: *"I don't think it's not plausible (he refers to THF). I wouldn't have intuitively guessed that something like this (THF) would happen if I am honest. But I would have taken an easier way and said it is a normal nucleophilic substitution, at this point, if I look at it that way I would not necessarily say it (THF) could be a by-product of what is being created. But the main product, the more I think about it, the more fascinating I think it is."*

Interviewer: *"Why?"*

Charlie: *"Because I really didn't think that this (formation of THF) was a possibility. Well, I didn't think about it at all, I even think that what was there before (nucleophilic substitution) is wrong, because I just saw that I have a nucleophile and a good leaving group and nucleophilic substitution, but I didn't see that you can build another nucleophile (alkoxide)."*

Charlie was one of the students who initially chose the diol (product card A) as the main product of the reaction of 4-chlorobutanol and hydroxide in task 2. When he received the alternative reaction product cards, he focused on THF as a reaction product. After he argued about both the alkoxide (product card D) and THF (product card E), he claimed both as the most plausible products of the reaction and explained that the alkoxide is the precursor of THF.

Among the students who applied evidence-based argumentation approaches, most did not provide a valid explanation for choosing THF as their most plausible product. Sonia, for instance, tried to build the mechanism for the formation of THF and could not find an acceptable explanation for her decision. As a result, the interviewer attempted to discuss the product THF in another way.

Interviewer: *"We can do it differently. When you see the product, would you describe the product as plausible in principle?"*

Sonia: *"Yes, but I don't know why. My feeling tells me that again, I don't know. It looks so right somehow."*

Interviewer: *"And are there any factors that help you determine why you think the product could be right?"*

Sonia: *"We once had a similar task in an exercise and I don't know, when I saw that, it kind of clicked in my head, somewhere in the back corner where I thought, I think that's it."*

Sonia later identified THF as the most plausible product of the reaction of 4-chlorobutanol and hydroxide. Nevertheless, she could not support her claim with any evidence or reasoning. She justified her claim based only on her feelings and on her memory of a similar task she completed sometime before. Like Sonia, no student with an evidence-based argumentation approach mentioned that the alkoxide (product card D) is the precursor for THF (product card E). Still, five out of nine students claimed THF as the most plausible product of the reaction with either little justification or without further justification. The following quote from Andy illustrates such a missing link from the precursor alkoxide to the reaction product THF. Andy built the diol (product card A) as a product in task 2 and chose the alkoxide as his first reaction product card to argue in task 3. He started to describe how the alkoxide was built.

Andy: *"So here, the hydroxyl group was simply deprotonated."*

Interviewer: *"How plausible do you think is this?"*

Andy: *"That's a good question. Actually, so it may happen. But a negative charge only on oxygen is very, I would say, unlikely."*

Interviewer: *"Why?" [...]*

Andy: *"So, that charge is on the oxygen and then is not somehow stabilised. That is a bit strange to me. Because, charge on an ester group would be kind of, I think, normal, because that's relatively stabilised by resonance. But only at the oxygen is a bit strange to me."*

Andy claimed incorrectly that the alkoxide is implausible as the reaction product. Since he could not explain why the charge

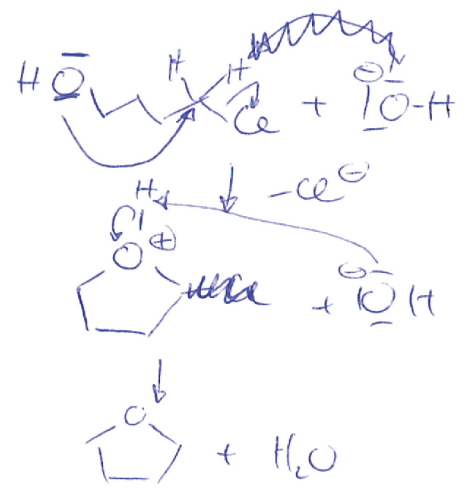


Fig. 9 Reaction mechanism for the formation of THF drawn by Andy.

on the oxygen atom seemed "strange" to him, he used an analogy with a molecule known to him, namely an ester, to support his claim. After Andy argued about the reaction product cards of the alkene (C) and the carbanion (B), he started to argue about THF (E) by drawing his idea of the reaction mechanism for the formation of THF (see Fig. 9).

Subsequently, the interviewer discussed the plausibility of the formation of THF (product card E) with Andy.

Interviewer: *"How plausible do you think is it?"*

Andy: *"That might even be more plausible than this (diol)."*

Interviewer: *"Okay. Why?"*

Andy: *"The O has lone pairs, a five-membered ring is not as strained as a four-membered-ring, for example. I would say that could be possible. So, I would even consider it relatively plausible."*

Both in Andy's drawing of the reaction mechanism and his explanation of the plausibility, he did not consider the alkoxide (product card D) as a precursor of THF formation. Andy proposed falsely that the hydroxyl group of 4-chlorobutanol attacks intramolecularly to form a five-membered ring that is deprotonated by the hydroxide ion in the last step. Since Andy needed several attempts to complete the reaction mechanism to his satisfaction, the interviewer asked him how he proceeded in setting up the reaction mechanism.

Andy: *"I noticed that the O has lone pairs. And then I thought, in order to reconstruct the five-membered-ring, it must attack here. That's where I first drew the Cl. But it would be a pentavalent carbon atom, which does not exist. So, the Cl has to get out somehow. And in the last step the H is taken by the OH. But only because the O is positive and it wants to become relatively neutral. And somehow furan* (Andy confounded THF and furan) *has to be the product."*

Interviewer: *"That means, you basically made sure that you knew the product and how to put it together?"*

Andy: *"Yes, somehow. Exactly."*

Comparable statements to those made by Andy, are well known in the literature. Bhattacharyya and Bodner (2005) showed that students tend to postulate a mechanism with the goal of forming a given product. However, in doing so, students are likely to build intermediates that are implausible.

In summary, Andy was not the only student who did not associate the alkoxide (product card D) with THF (product card E). Nevertheless, while most students changed their initial claim, the rationale for the change varied greatly. In particular, students who used a reasoning-based argumentation approach identified the alkoxide as the precursor of THF and provided a rationale for it. Although five out of nine students who applied an evidence-based argumentation approach also chose THF as the most plausible product, they were not able to (sufficiently) justify their decision.

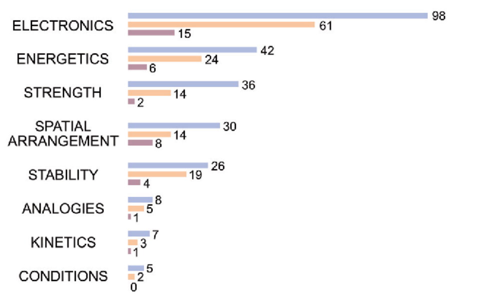


Fig. 10 Absolute number of codes attributed to reasoning statements in the three argumentation approaches: reasoning-based argumentation approach (blue), based-on-both argumentation approach (orange), and evidence-based argumentation approach (purple).

How do students use reasoning to justify their arguments while assessing the plausibility of alternative reaction products?

As described in the data analysis section, the analysis of students' reasoning statements resulted in eight data-based reasoning categories, which are shown in Fig. 11. Examples of the four most frequently used reasoning categories illustrate which type of statements were assigned to the categories. The most frequently used category was *Electronics*, which refers to factors such as polarisation and electronegativity. Students' statements coded in this category include "due to the positive inductive effect, which is an electron-pushing effect, more electrons would be distributed" (Ronnie, based-on-both argumentation approach) and "negative charge next to negative charge is unfavourable because there is a high excess of electrons" (Haley, evidence-based argumentation approach). The *Energetics* category covers energetic aspects such as bonds and energies. One example is the correct statement given by Pepper (reasoning-based argumentation approach), to justify the plausibility of THF (product card E): "that is because the entropy increases. The degeneracy in the system increases, since the energy can be distributed over more molecules." Additional examples are the statements made by many students when talking about the plausibility of the carbanion (reaction product card B) or the alkene (reaction product card C) related to the high strength of the C–H bond because a lot of energy is required to cleave the bond. The *Spatial Arrangement* category includes statements that refer to aspects of sterics, distance, or size. Stella (based-on-both argumentation approach) justified her evidence incorrectly that chloride is a good leaving group as follows: "that chloride, as an ion, is smaller and can move easily compared to the hydroxide ion." On the other hand, Sal (evidence-based argumentation approach) justified that the hydroxyl group is less reactive as follows: "because the carbon chain is quite long." The last of the four most frequently used categories was *Strength*, which includes the quality or strength of leaving groups or acids and bases. Statements in this category include several students' justification of good leaving groups by referring to the base strength and the justification for acids and bases based on the (estimated) pK_a values, as cited by over a quarter of the students.

Clear differences in the numbers of reasoning statements between argumentation approaches can be seen in Fig. 11. The reasoning-based argumentation approach (blue) was associated with the use of considerably more reasoning statements compared to the based-on-both argumentation approach (orange) and the evidence-based argumentation approach (purple), although a distinction is also visible between the latter two approaches. However, the percentages of reasoning statements in the different reasoning categories were similar across the three argumentation approaches (see Appendix 3). For example, the *Electronics* category accounted for 39% for the reasoning statements for the reasoning-based argumentation approach, 43% for the based-on-both argumentation approach, and 40% for the evidence-based argumentation approach. The percentages were also similar for the *Energetics* category (17% for the reasoning-based argumentation approach, 17% for the based-on-both argumentation approach, 16% for the evidence-based argumentation approach) and the *Kinetics* category (3% for the reasoning-based argumentation approach, 2% for the based-on-both argumentation approach, and 3% for the evidence-based argumentation approach). The students not only differed in their number of reasoning statements, but also in the way they justified their statements. For example, Cameron (reasoning-based argumentation approach) and Gloria (evidence-based argumentation approach) both mentioned the correct evidence that chloride is a good leaving group. However, these two students differed in how they justified this evidence.

Cameron: *"Chloride is a good leaving group because the chlorocarbon bond is not very stable, unlike leaving groups such as OH^-. Then it can be stabilised in an aqueous solution.* [. . .] *There is also a high electronegativity. It's also very electron-withdrawing, chlorine as an atom."*

Interviewer: *"Could you explain to me why chloride is a good leaving group and hydroxide is a weak leaving group?"*

Cameron: "[. . .] *So theoretically it should be that the electron that is added in chlorine is stored in an orbital, which is energetically higher than hydroxide. But I'm not really getting anywhere with it (laughs). So, I would say that this is not the cause because of the stability of the two ions, but because of the stability of the bond. That the splitting, here with chlorine and then there is another bond with carbon, that the splitting, if I were to occupy that as well, of the orbitals is lower with chlorine than with oxygen. And therefore, if this energy gain, through the bond or the energy loss through the cleavage of the bond, is then lower with chlorine than with OH."*

Cameron provided multiple reasoning statements for his evidence, including statements in different reasoning categories. In the *Energetics* category, for example, he stated, that the energy loss or energy gain in the case of the cleavage of the C–Cl bond is energetically more favourable than the cleavage of the C–O bond. Furthermore, Cameron stated that the orbital of the chlorine atom involved in the reaction is higher in energy than the hydroxide orbital. In general, few students included orbitals in their reasoning. Cameron also mentioned the stabilisation of the chloride ion in water (*Stability*) and the electronegativity of chlorine (*Energetics*) as factors that make chloride a better leaving group than hydroxide. Gloria also cited the fact that chloride is a good leaving group as evidence.

Gloria: *"Halogens are generally a good leaving group. That's why I suspect a nucleophilic substitution."*

Interviewer: *"Can you tell me why halogens are good leaving groups?"*

Gloria: *"Because Cl^- in particular is relatively stable and obeys the octet rule."*

Gloria made two reasoning statements. On one hand, she described chloride as a relatively stable ion (*Stability*). On the other hand, she mentioned the obeyed octet rule (*Electronics*) as a factor to justify her evidence. In contrast to Cameron, Gloria did not provide any further information on the stabilisation of the chloride ion, even when asked. In addition, she did not draw a comparison to the hydroxide ion, which competes with the chloride ion as a possible leaving group. When prompted, she could not provide any reasoning for why hydroxide is a worse leaving group compared to the chloride. This is in line with the analysis of Popova and Bretz (2018b), who found that students often have problems when asked to justify why a certain leaving group is good or bad. Directly comparing the statements of Cameron and Gloria regarding the same evidence indicates that both the number of reasoning statements and the depth of the reasoning statements are important for substantiating an argument.

Conclusion and implications

The building of arguments in the field of organic chemistry is a central ability for students to make well-grounded decisions. Arguing about multiple reaction pathways requires the consideration of multiple chemical variables that influence the pathway of a reaction. Thus, principles that are often considered separately by students must be brought together in a meaningful way to weigh them against each other (Watts *et al.*, 2021). Analytical thinking can be encouraged by providing students with the opportunity to include different aspects in their decision-making process, as has been reported for tasks based on alternative reaction pathways (Lieber and Graulich, 2020). In this study, we investigated how students argue when assessing the plausibility of alternative reaction pathways. We used a simplified version of Toulmin's argumentation model, which can be considered as the key components of an argument (Toulmin, 2003; Osborne and Patterson, 2011).

In analysing students' argumentation patterns, we specifically asked the following three questions. (1) To what extent do students make use of evidence and reasoning in their argumentation (*i.e.*, argumentation approach)? (2) How do students with different argumentation approaches rationalise changes in their initial claims? (3) How do students use reasoning to justify their arguments while assessing the plausibility of alternative reaction pathways?

The analysis revealed that students can be categorised into three argumentation approaches (reasoning-based, based-on-both, and evidence-based) based on the frequencies of evidence and reasoning statements and their ratio. There are different ways to form arguments, and it should be emphasised that there is no ideal argumentation approach. Instead, the numbers of evidence and reasoning statements were used in this study as an indicator of the potential quality of argumentation. A large number of evidence statements suggests that the student has recognised and included many different aspects in the decision-making process and considered various chemical variables influencing the reaction process. A large number of evidence statements can thus increase the quality and breadth of an argument since the course of a chemical reaction is influenced by a variety of factors. Recognising different influencing factors can also enhance the quality of the weighing process if the evidence is supported by reasoning. The number of reasoning statements is an indicator of how precise or elaborated the justification of evidence is. The more precise the justification, the greater the depth of the built argument. Proper justifying statements increase the quality of an argument by creating a stronger link between the components of the argument and the required conceptual understanding, which is a crucial aspect of the quality of an argument (Sandoval and Millwood, 2005; Choi *et al.*, 2013). A well-grounded foundation of evidence and reasoning gives weight to an argument and should be an integral part of the classroom culture (Walker *et al.*, 2012; Towns *et al.*, 2019).

In addition to examining students' use of evidence and reasoning, we assessed whether students changed their initial claims and their ability to rationalise these changes. Solving tasks that require the student to build multiple arguments can lead to more thorough decision-making, possibly enabling the student to justifiably defend a claim while also providing the opportunity to revise the claim. Based on the analysis of students' argumentation patterns, 21 out of the 29 students in this study questioned their decisions and revised their claims during the argumentation process. Still, it seems that some students reached the limits of their conceptual understanding when rationalising their decisions; these students were unable to elaborate on their decisions beyond a certain individual point, even when specifically prompted. This issue was particularly noticeable for students who used the evidence-based argumentation approach, as these students were often unable to justify their claims. Luo *et al.* (2020) and Walker *et al.* (2019) found that students struggled to change their initial claims and thus stuck to their often erroneous claims. In the present study, the students were likely to question their initial claims after being specifically prompted to defend or revise them. Luo *et al.* (2020) analysed written arguments, whereas our study was based on interviews, which may lead students to be more willing to change their claim after prompting.

The findings of this study have some implications for teaching. In accordance with former studies, when introducing the building of arguments, teachers should ensure that students do not stop at supporting the claim with evidence, but also justify the evidence with reasoning (Bodé *et al.*, 2019; Crandell *et al.*, 2020). This part of argumentation is where students are likely to have the most problems; thus, it is important that teachers value justification through reasoning, even if it may be erroneous, to motivate students to build complete CER argumentation patterns. To help students learn to support and justify claims, argumentation patterns similar to concept maps can be used in teaching. As shown in Fig. 4 and Appendix 2, these argumentation structures can be used to create "argumentation trees" in the form of CER trees. This might help learners and teachers visualise both the depth and breadth of arguments and highlight gaps.

Furthermore, building arguments for alternative reaction pathways can be helpful as it helps students recognise that chemical reactions in the field of organic chemistry can lead to by-products (Popova and Bretz, 2018a) or may not form the desired product due to the choice of reaction conditions and reagents. In this study, we extended a common task format (*i.e.*, comparing two reaction paths or reaction products) to the consideration of five alternative reaction pathways.

Students in this study were divided into three argumentation approaches based on the number of reasoning and evidence statements and their ratio. Based on the individual arguments formed by the students, the students need further support to develop strong arguments. Compared to the other two argumentation approaches, students who applied a reasoning-based argumentation approach provided a larger number of evidence and reasoning statements. To encourage this level of argumentation in future tasks, students should be supported in justifying their claims (*e.g.*, by being trained to use several pieces of evidence for each claim and to justify each piece of evidence with several reasoning statements). In addition to the breadth of an argument, the depth of the argument is also of interest. In this regard, students could be provided with conceptual support to improve the quality of their evidence and reasoning statements. Additional support could focus on the weighing process through which students consider their formed evidence and reasoning statements. Students using the based-on-both argumentation approach supported their claims with multiple pieces of evidence and reasoning. However, the ratio of evidence to reasoning statements was close to 1, meaning that each evidence statement was justified on average with one reasoning statement. Thus, students in this group could receive targeted reasoning training to improve the breadth of their arguments and increase the number of reasoning statements for each piece of evidence. Students who used the evidence-based argumentation approach hardly provided any evidence or reasoning statements on their own and often failed to justify them, even after targeted prompting. On the one hand, this may be due to the fact that the students were conceptually unable to support their statements. Therefore, these students would benefit from additional conceptual support (*e.g.*, providing information that they should include in their arguments). On the other hand, this group of students might benefit from examples highlighting the difference between evidence and reasoning statements or additional training on structuring an argument.

Scaffolds such as those used by Luo *et al.* (2020) and Caspari and Graulich (2019) to slow down students' decision-making process (Caspari and Graulich, 2019) have proven to be helpful and could be adapted to help students build arguments for multiple alternative pathways. Students could be engaged to collect evidence as a first step and then be prompted to purposefully provide reasoning linked to the evidence. However, when developing evidence and reasoning to assess the plausibility of alternative reaction pathways, the quality of the evidence can vary, leading students to struggle in justifying their claims. Some students may profit from conceptual support in addition to support in building of arguments. An ongoing study is investigating whether students' argumentation patterns in organic chemistry can be supported with an adaptive scaffold that targets the student's level of conceptual knowledge and argumentation skills (*i.e.*, the use of evidence and reasoning).

Limitations

The conclusions drawn should be considered with caution given certain inherent limitations of this study. First, the students in this qualitative interview study voluntarily participated and may have been motivated to participate by different reasons. Nevertheless, 29 out of 40 students in the course agreed to participate in the study, which allowed this study to cover the spectrum of student performance in the course. Additionally, the interview format may also lead students to engage more intensively with the topic than when conducting tasks on their own. Second, the consideration of alternative reaction pathways may be a new concept for the students in this course, leading to cognitive overload. Moreover, considering the students' previous course content, students were not used to building arguments or explaining the decisions that led to their solutions in detail. This may have resulted in students not expressing all their thoughts aloud, leading to individual cases of shortened argumentation being presented. The interviewer attempted to counteract this by asking specific questions, which were rephrased several times if necessary, to get each student to present their thoughts as precisely as possible. Third, the correctness of students' statements was not considered since we wanted to assess the students' thought processes in their entirety. Thus, erroneous statements that the student was not aware of were also considered. However, to determine what support is needed, it is essential to consider the complete thought process and not only erroneous or correct components. Lastly, our focus on the ratio of reasoning to evidence statements simplified the students' argumentation processes and might not have captured subtle differences.

Conflicts of interest

There are no conflicts to declare.

Appendix 1: overview of the distribution of students' evidence and reasoning statements

Fig. 11 shows the numbers of students' evidence and reasoning statements for each alternative reaction product card

Name	DIOL		CARBANION		ALKENE		ALKOXIDE		THF		TOTAL		RATIO
	Evidence	Reasoning	Evidence	Reasoning	Evidence	Reasoning	Evidence	Reasoning	Evidence	Reasoning	Evidence	Reasoning	Reasoning/Evidence
Alex	4	4	5	4	5	6	1	2	3	4	18	20	1,11
Lily	4	5	3	1	4	10	1	2	3	3	15	21	1,40
Jay	3	5	5	9	4	8	3	6	4	5	19	33	1,74
Pepper	4	7	2	1	1	1	3	4	5	8	15	21	1,40
Luke	6	6	2	1	7	8	4	6	5	4	24	25	1,04
Cameron	5	6	3	3	5	6	2	2	7	7	22	24	1,09
Charlie	2	6	3	3	3	6	3	1	4	7	15	23	1,53
Phil	5	10	2	8	2	2	2	4	3	4	14	28	2,00
Claire	5	7	6	5	3	3	3	3	3	5	20	23	1,15
Reuben	3	7	4	7	3	7	5	6	2	4	17	31	1,82
Ben	5	9	2	1	0	0	3	1	2	2	12	13	1,08
Joe	3	4	3	4	2	2	2	3	0	0	10	13	1,30
Stella	2	4	2	5	2	5	1	1	1	1	8	16	2,00
Manny	4	6	5	6	5	3	2	1	3	0	19	16	0,84
Mitchell	4	4	2	2	2	3	2	2	4	5	14	16	1,14
Dylan	3	2	4	2	2	2	2	1	4	5	15	12	0,80
Gil	2	5	3	4	3	1	2	2	3	5	13	17	1,31
Sydney	2	5	1	2	3	3	3	2	3	2	12	14	1,17
Rhonda	3	5	2	2	1	3	1	0	2	2	9	12	1,33
Ronnie	3	3	3	3	2	1	3	4	1	2	12	13	1,08
Amber	3	3	2	1	0	0	1	0	1	0	7	4	0,57
Haley	1	0	1	1	4	1	2	0	3	1	11	3	0,27
Gloria	4	3	2	1	1	0	0	0	1	1	8	5	0,63
Pam	2	2	3	2	1	0	2	1	2	1	10	6	0,60
Sonia	1	1	1	2	2	3	1	0	1	0	6	6	1,00
Sal	2	4	2	0	1	0	2	1	1	1	8	6	0,75
Andy	2	1	2	0	0	0	1	2	2	0	7	3	0,43
Beth	2	2	3	1	0	0	2	1	0	0	7	4	0,57
Frank	1	0	1	1	1	0	1	0	3	2	7	3	0,43

Fig. 11 Overview of the distribution of each student's evidence and reasoning statements for each of the five alternative reaction product cards, the total of all evidence and reasoning statements, and the ratio of reasoning statements to evidence statements.

(diol, alkene, carbanion, alkoxide, and THF), the total of all evidence and reasoning statements for each student, and the ratio of reasoning statements to evidence statements (on which the classification of students into the three argumentation approaches was based). The three argumentation approaches are colour-coded as follows in Fig. 11: the reasoning-based argumentation approach is highlighted in blue, the based-on-both argumentation approach is highlighted in orange, and the evidence-based argumentation approach is highlighted in purple.

Appendix 2: example total argumentation patterns for the three argumentation approaches

Three example argumentation patterns for all product cards for each argumentation approach (reasoning-based, based-on-both, and evidence-based) are shown in the Fig. 12–14 below. Fig. 12 illustrates the argumentation patterns for Reuben (reasoning-based argumentation approach), Fig. 13 illustrates

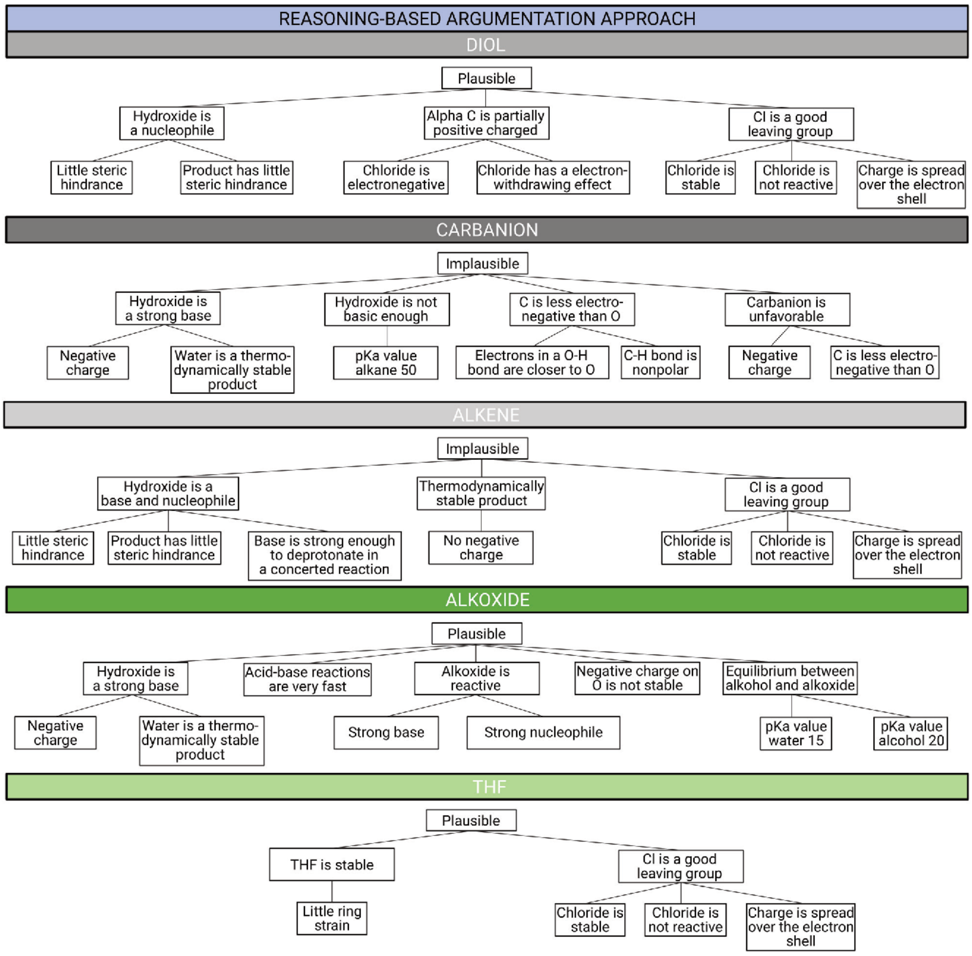


Fig. 12 Example argumentation pattern (CER structure) for all product cards of the reasoning-based argumentation approach from Reuben.

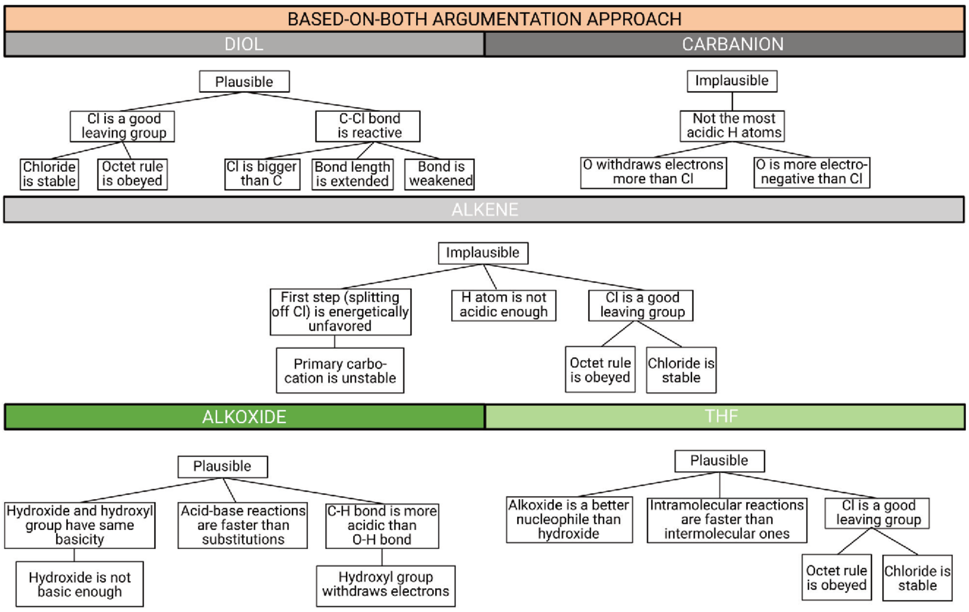


Fig. 13 Example argumentation pattern (CER structure) for all product cards of the based-on-both argumentation approach from Sydney.

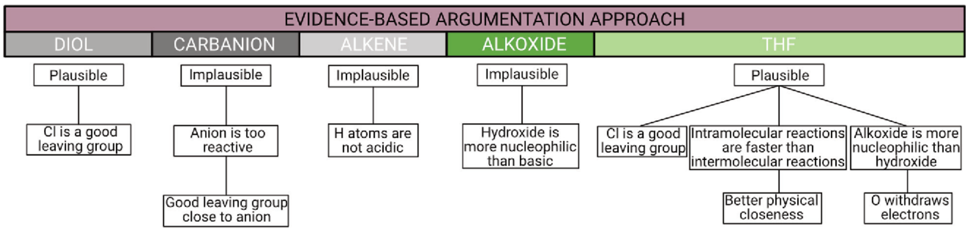


Fig. 14 Example argumentation patterns (CER structure) for all product cards of the evidence-based argumentation approach from Frank.

those for Sydney (based-on-both argumentation approach), and Fig. 14 shows those of Frank (evidence-based argumentation approach). All students built arguments for the plausibility of the five alternative reaction products. The arguments were analysed using the CER structure.

Appendix 3: distribution of statements among the different reasoning categories

The total number of reasoning statements varied greatly among the three argumentation approaches (Fig. 10). Nevertheless, the

Table 2 Percentage of reasoning statements in different categories for the three argumentation approaches (approach 1 = reasoning-based argumentation approach, approach 2 = based-on-both argumentation approach, approach 3 = evidence-based argumentation approach)

	Electronics (%)	Energetics (%)	Kinetics (%)	Spatial arrangement (%)	Analogies (%)	Strength (%)	Conditions (%)
Approach 1	39	17	3	12	3	14	2
Approach 2	43	17	2	10	4	10	1
Approach 3	40	16	3	22	3	5	0

percentage of reasoning statements in the different categories were similar among the three approaches. Table 2 illustrates the percentages of reasoning statements in each category for the three argumentation approaches.

Acknowledgements

This publication is part of the first author's doctoral (Dr rer. nat.) thesis at the Faculty of Biology and Chemistry, Justus-Liebig-University Giessen, Germany. We are thankful for all students willing to participate in the study. Moreover, we thank all members of the Graulich group for fruitful discussions, Axel Langner for his support with graphics, and especially Krenare Ibraj. Leonie Lieber thanks the Verband der Chemischen Industrie (German Chemical Industry Association) for supporting her with the Kekulé Fellowship.

References

Abi-El-Mona I. and Abd-El-Khalick F., (2011), Perceptions of the nature and 'goodness' of argument among college students, science teachers, and scientists, *Int. J. Sci. Educ.*, **33**(4), 573–605.

Akkuzu N. and Uyulgan M. A., (2016), An epistemological inquiry into organic chemistry education: Exploration of undergraduate students' conceptual understanding of functional groups, *Chem. Educ. Res. Pract.*, **17**(1), 36–57.

Anzovino M. E. and Bretz S. L., (2015), Organic chemistry students' ideas about nucleophiles and electrophiles: The role of charges and mechanisms, *Chem. Educ. Res. Pract.*, **16**(4), 797–810.

Bacon F., (1878), *Novum organum*, Clarendon Press.

Becker N., Rasmussen C., Sweeney G., Wawro M., Towns M. and Cole R., (2013), Reasoning using particulate nature of matter: An example of a sociochemical norm in a university-level physical chemistry class, *Chem. Educ. Res. Pract.*, **14**(1), 81–94.

Becker N., Noyes K. and Cooper M., (2016), Characterizing students' mechanistic reasoning about London dispersion forces, *J. Chem. Educ.*, **93**(10), 1713–1724.

Bernard H. R. and Bernard H. R., (2013), *Social research methods: Qualitative and quantitative approaches*, Sage.

Bhattacharyya G. and Bodner G. M., (2005), "It gets me to the product": How students propose organic mechanisms, *J. Chem. Educ.*, **82** (9), 1402–1407.

Bodé N. E., Deng J. M. and Flynn A. B., (2019), Getting past the rules and to the WHY: Causal mechanistic arguments when judging the plausibility of organic reaction mechanisms, *J. Chem. Educ.*, **96**, 1068–1082.

Cartrette D. P. and Mayo P. M., (2011), Students' understanding of acids/bases in organic chemistry contexts, *Chem. Educ. Res. Pract.*, **12**(1), 29–39.

Caspari I. and Graulich N., (2019), Scaffolding the structure of organic chemistry students' multivariate comparative mechanistic reasoning, *Int. J. Phys. Chem. Educ.*, **11**(2), 31–43.

Caspari I., Kranz D. and Graulich N., (2018), Resolving the complexity of organic chemistry students' reasoning through the lens of a mechanistic framework, *Chem. Educ. Res. Pract.*, **19** (4), 1117–1141.

Chinn C. A. and Brewer W. F., (2001), Models of data: A theory of how people evaluate data, *Cogn. Instruct.*, **19**(3), 323–393.

Choi A., Hand B. and Greenbowe T., (2013), Students' written arguments in general chemistry laboratory investigations, *Res. Sci. Educ.*, **43**(5), 1763–1783.

Cooper M. M., (2015), Why ask why?, *J. Chem. Educ.*, **92**(8), 1273–1279.

Cooper M. M., Kouyoumdjian H. and Underwood S. M., (2016), Investigating students' reasoning about acid–base reactions, *J. Chem. Educ.*, **93**(10), 1703–1712.

Crandell O. M., Lockhart M. A. and Cooper M. M., (2020), Arrows on the page are not a good gauge: Evidence for the importance of causal mechanistic explanations about nucleophilic substitution in organic chemistry, *J. Chem. Educ.*, **97**(2), 313–327.

Cruz-Ramirez de Arellano D. and Towns M. H., (2014), Students' understanding of alkyl halide reactions in undergraduate organic chemistry, *Chem. Educ. Res. Pract.*, **15**(4), 501–515.

DeFever R. S., Bruce H. and Bhattacharyya G., (2015), Mental rolodexing: Senior chemistry majors' understanding of chemical and physical properties, *J. Chem. Educ.*, **92**(3), 415–426.

Deng J. M. and Flynn A. B., (2021), Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items, *Chem. Educ. Res. Pract.*

Driver R., Newton P. and Osborne J., (2000), Establishing the norms of scientific argumentation in classrooms, *Sci. Educ.*, **84**(3), 287–312.

Erduran S., (2019), *Argumentation in Chemistry Education*, Royal Society of Chemistry, pp. 1–10.

Franck G., (2012), Modern science: A case of collective intelligence? On the role of thought economy and gratifying attention in knowledge production, *Angew. Chem., Int. Ed.*, **51**(29), 7088–7092.

Galloway K. R., Leung M. W. and Flynn A. B., (2019), Patterns of reactions: A card sort task to investigate students' organization of organic chemistry reactions, *Chem. Educ. Res. Pract.*, **20**(1), 30–52.

Graulich N., Hedtrich S. and Harzenetter R., (2019), Explicit versus implicit similarity – Exploring relational conceptual understanding in organic chemistry, *Chem. Educ. Res. Pract.*, **20**(4), 924–936.

Grove N. P., Cooper M. M. and Cox E. L., (2012), Does mechanistic thinking improve student success in organic chemistry?, *J. Chem. Educ.*, **89**(7), 850–853.

Hand B. and Choi A., (2010), Examining the impact of student use of multiple modal representations in constructing arguments in organic chemistry laboratory classes, *Res. Sci. Educ.*, **40**(1), 29–44.

Hogan K. and Maglienti M., (2001), Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions, *J. Res. Sci. Teach.*, **38**(6), 663–687.

Hosbein K. N., Lower M. A. and Walker J. P., (2021), Tracking student argumentation skills across general chemistry through argument-driven inquiry using the assessment of scientific argumentation in the classroom observation protocol, *J. Chem. Educ.*

Kraft A., Strickland A. M. and Bhattacharyya G., (2010), Reasonable reasoning: multi-variate problem-solving in organic chemistry, *Chem. Educ. Res. Pract.*, **11** (4), 281–292.

Kuhn D., (1991), *The Skills of Argument*, Cambridge University Press.

Lazarou D. and Erduran S., (2021), "Evaluate what I was taught, not what you expected me to know": Evaluating students' arguments based on science teachers' adaptations to toulmin's argument pattern, *J. Sci. Teach. Educ.*, **32**(3), 306–324.

Lieber L. and Graulich N., (2020), Thinking in alternatives-A task design for challenging students' problem-solving approaches in organic chemistry, *J. Chem. Educ.*, **97** (10), 3731–3738.

Loibl K. and Leuders T., (2019), How to make failure productive: Fostering learning from errors through elaboration prompts, *Learn. Instr.*, **62**, 1–10.

Luo X. L., Wei B., Shi M. and Xiao X., (2020), Exploring the impact of the reasoning flow scaffold (RFS) on students' scientific argumentation: based on the structure of observed learning outcomes (SOLO) taxonomy, *Chem. Educ. Res. Pract.*, **21** (4), 1083–1094.

McNeill K. L. and Krajcik J., (2012), *Book study facilitator's guide: Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*, Pearson Allyn & Bacon, New York.

McNeill K. L., Lizotte D. J., Krajcik J. and Marx R. W., (2006), Supporting students' construction of scientific explanations by fading scaffolds in instructional materials, *J. Learn. Sci*, **15** (2), 153–191.

Osborne J. F. and Patterson A., (2011), Scientific argument and explanation: A necessary distinction?, *Sci. Educ.*, **95** (4), 627–638.

Petritis S. J., Kelley C. and Talanquer V., (2021), Exploring the impact of the framing of a laboratory experiment on the nature of student argumentation, *Chem. Educ. Res. Pract.*, **22** (1), 105–121.

Platt J. R., (1964), Strong Inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others, *Science*, **146** (3642), 347–353.

Popova M. and Bretz S. L., (2018a), "It's only the major product that we care about in organic chemistry": An analysis of students' annotations of reaction coordinate diagrams, *J. Chem. Educ.*, **95** (7), 1086–1093.

Popova M. and Bretz S. L., (2018b), Organic chemistry students' understandings of what makes a good leaving group, *J. Chem. Educ.*, **95** (7), 1094–1101.

Sadler T. D., (2004), Informal reasoning regarding socioscientific issues: A critical review of research, *J. Res. Sci. Teach.*, **41** (5), 513–536.

Sandoval W. A. and Millwood K. A., (2005), The quality of students' use of evidence in written scientific explanations, *Cogn. Instruct.*, **23** (1), 23–55.

Talanquer V., (2006), Commonsense chemistry: A model for understanding students' alternative conceptions, *J. Chem. Educ.*, **83** (5), 811–816.

Talanquer V., (2018), Importance of understanding fundamental chemical mechanisms, *J. Chem. Educ.*, **95**(11), 1905–1911.

Toulmin S. E., (2003), *The Uses of Argument, Updated Version*, Cambridge University Press.

Towns M. H., Cole R. S., Moon A. C. and Stanford C., (2019), *Argumentation in Chemistry Education*, Royal Society of Chemistry, pp. 247–274.

Turro N. J., (1986), Geometric and topological thinking in organic-chemistry, *Angew. Chem., Int. Ed. Engl.*, **25**(10), 882–901.

Van Eemeren F. H., Garssen B., Krabbe E. C. W., Francisca Snoeck Henkemans A., Verheij B. and Wagemans J. H. M., (2014), *Handbook of Argumentation Theory*, Dordrecht: Springer, pp. 203–256.

Walker J. P., Sampson V., Grooms J., Anderson B. and Zimmerman C. O., (2012), Argument-driven inquiry in undergraduate chemistry labs: The impact on students' conceptual understanding, argument skills, and attitudes toward science, *J. Coll. Sci. Teach.*, **41**(4), 74–81.

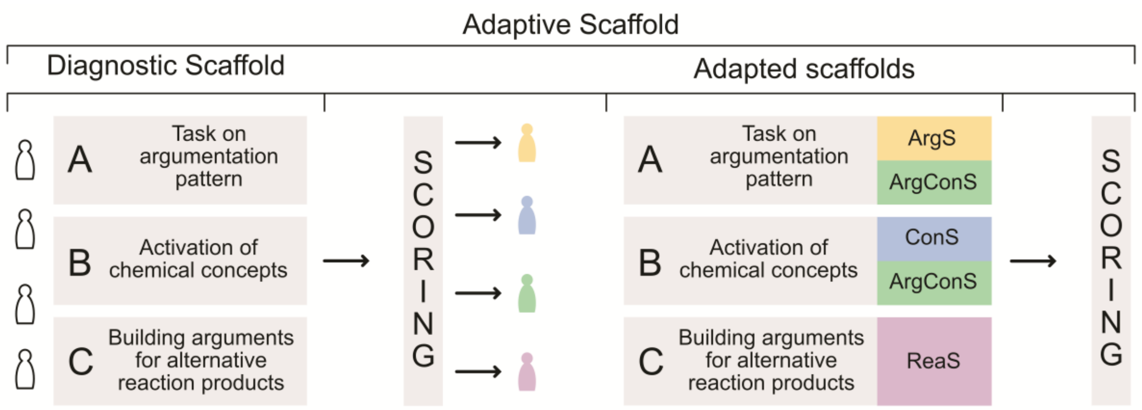
Walker J. P., Van Duzor A. G. and Lower M. A., (2019), Facilitating argumentation in the laboratory: The challenges of claim change and justification by theory, *J. Chem. Educ.*, **96** (3), 435–444.

Watts F. M., Schmidt-McCormack J. A., Wilhelm C. A., Karlin A., Sattar A., Thompson B. C., Gere A. R. and Shultz G. V., (2020), What students write about when students write about mechanisms: Analysis of features present in students' written descriptions of an organic reaction mechanism, *Chem. Educ. Res. Pract.*, **21** (4), 1148–1172.

Watts F. M., Zaimi I., Kranz D., Graulich N. and Shultz G. V., (2021), Investigating students' reasoning over time for case comparisons of acyl transfer reaction mechanisms, *Chem. Educ. Res. Pract.*, **22** (2), 364–381.

4 Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns

In Chapter 4, the original paper is represented. It is reproduced from Lieber, Ibraj, Caspari-Gnann and Graulich (2022) with permission of the Royal Society of Chemistry.



Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns. *Chemistry Education Research and Practice*, advance article, DOI: 10.1039/d2rp00016d.

Chemistry Education Research and Practice

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/d2rp00016d

Closing the gap of organic chemistry students' performance with an adaptive scaffold for argumentation patterns

Leonie Sabine Lieber,[a] Krenare Ibraj,[a] Ira Caspari-Gnann[b] and Nicole Graulich*[a]

Building reasonable scientific arguments is a fundamental skill students need to participate in scientific discussions. In organic chemistry, students' argumentation and reasoning skills on reaction mechanisms are described as indicators of success. However, students often experience challenges with how to structure their arguments, use scientific principles appropriately and engage in multivariate, instead of one-reason decision-making. Since every student experiences their individual challenges with a multitude of expectations, we hypothesise that students would benefit from scaffolding that is adapted to their needs. In the present study, we investigated how 64 chemistry students interacted with an adaptive scaffold that offered different ways of support based on students' strengths and limitations with structural and conceptual aspects that are needed to build a scientific argument in organic chemistry. Based on the students' performance in a diagnostic scaffold in which they were asked to judge the plausibility of alternative organic reaction pathways by building arguments, the students were assigned to one of four support groups that received a scaffold adapted to their respective needs. Comparing students' performance in the diagnostic and adapted scaffolds allows us to determine quantitatively (1) to what extent the adaptive scaffold closes the gap in students' performance and (2) whether an adaptive scaffold improves the students' performance in their respective area of support (argumentation and/or concept knowledge). The results of this study indicate that the adaptive scaffold can adaptively advance organic chemistry students' argumentation patterns.

Received 10th January 2022,
Accepted 30th May 2022

DOI: 10.1039/d2rp00016d

rsc.li/cerp

Introduction

The overall goal of chemistry education is to enable students to participate in scientific discourse (McGinn and Roth, 1999; Stowe *et al.*, 2021) by providing them with the tools to think critically, explain phenomena, and build reasoned arguments (Driver *et al.*, 1994). However, the way how we want students to build arguments often differs from how students actually build, use, and understand argumentation (McNeill and Krajcik, 2012). The most common observation is that students experience difficulties supporting their claim with evidence and justifying it with reasoning when building arguments (Deng and Flynn, 2021; Petritis *et al.*, 2021; Lieber and Graulich, 2022), as students often do not know what counts as evidence or reasoning (Sadler, 2004) or tend to limit their argumentation to evidence and overlook reasoning. Moreover, students experience challenges with using scientific rules and principles when connecting evidence and reasoning (McNeill and Krajcik, 2012; Walker *et al.*, 2019). As a result, students may rely on personal views such as gut feelings (Hogan and Maglienti, 2001) as it avoids uncertainties with scientific principles, among other reasons. In addition to the reported challenges with structuring an argument, applying conceptual understanding when building arguments can also be challenging. Difficulties with conceptual knowledge might hinder building chemically knowledgeable and multivariate arguments (Cruz-Ramirez de Arellano and Towns, 2014; Moon *et al.*, 2016; Pabuccu and Erduran, 2017; Deng and Flynn, 2021; Lieber and Graulich, 2022). Thereby, students make use of single concepts to justify their claim, even when it is necessary to include multiple concepts in their decision-making (McNeill *et al.*, 2006).

Ongoing research on student argumentation clearly shows that the reported difficulties are either caused by (1) missing knowledge on how to structure an argument or missing activation of that knowledge for the problem at hand or (2) missing conceptual understanding or its application required for the argument or (3) both reasons. Given that prior knowledge has a major impact on students' performance, students need to

[a] *Justus-Liebig University Giessen, Institute of Chemistry Education, Heinrich-Buff-Ring 17, 35392 Giessen, Germany. E-mail: Nicole.graulich@dc.jlug.de*
[b] *Department of Chemistry, Tufts University, 62 Talbot Ave, Medford, MA 02155, USA*

receive adapted support to build upon their strengths and limitations with argumentation, and the level of conceptual understanding they bring into the classroom (Chen, 2014). Scaffolds tailored to students' needs may support them adaptively to solve a task on their own and by purposefully guiding and slowing down specific aspects of the argumentation process. It may, thus, support the students to direct their focus to the expected structure of an argument or to consider conceptual understanding that they may not have activated without a scaffold (Wood *et al.*, 1976). Especially when building arguments for multivariate mechanistic tasks, slowing down the reasoning process with a scaffold has been shown to assist learners in first collecting multiple relevant chemical concepts and weighing them afterward before making a decision (Caspari and Graulich, 2019; Flynn, 2021; Watts *et al.*, 2021). McNeill *et al.* (2006) emphasised that scaffolding as a flexible process should not be rigid. Instead, scaffolding should be adjusted to students' needs.

Previous research on argumentation in chemistry demonstrated that (1) students experience challenges with building sound arguments, (2) students experience difficulties using appropriate scientific principles, and (3) scaffolds are powerful tools to address students' needs. However, adaptive scaffolds designed to close this performance gap are still limited (Chen, 2014).

In this study, students received scaffolded training consisting of two consecutive parts (*i.e.*, two different data points) which we refer to as an adaptive scaffold. In the first part of the adaptive scaffold, which we refer to as 'diagnostic scaffold' students received practice and support in building arguments and using concept knowledge. Students' answers in this diagnostic scaffold served as a diagnosis for the second part of the adaptive scaffold in which students received one of four scaffolds. These adapted scaffolds addressed the previously mentioned difficulties, resulting in four adapted scaffolds for argumentation patterns, the use of concept knowledge, argumentation patterns and the use of concept knowledge, and students with no apparent difficulties.

Therefore, this quantitative study reports on the effectiveness of the adaptive scaffold (*i.e.*, a combination of a diagnostic scaffold and four adapted scaffolds), designed as an online learning environment to adaptively scaffold students based on their performance of building arguments for alternative organic reaction pathways. In this manner, we investigated the extent to which the adaptive scaffold closes the gap in organic chemistry students' performance and whether the adaptive scaffold improved students' performance.

Theoretical framework

Claim-evidence-reasoning (CER) argumentation model

Toulmin (2003) postulated that an argument consists of six elements: a claim that is supported by data, a warrant that bridges the gap between data and claim, and a backing for the warrant. The argument also includes a qualifier and a rebuttal that weaken the claim by questioning its necessity and whether there are exceptions to the rule (Toulmin, 2003). However, the model is often too complex to introduce an argument structure to students, *e.g.*, distinguishing between data, warrant, and backing is challenging (Erduran, 2007). Moreover, both students and teachers reported that some elements of Toulmin's argumentation model cause ambiguity (Lazarou and Erduran, 2020). Toulmin already stated that not every argument necessarily has to consist of all six components. Instead, he emphasised that the core of an argument only includes claim, data, and warrant (Toulmin, 2003). For this reason, a simplified argumentation model, the CER model, is frequently used, which consists of three components: claim, evidence and reasoning (McNeill *et al.*, 2006; McNeill and Krajcik, 2012). Both argumentation models have in common that a claim is supported by evidence and is justified with reasoning. However, the CER claim is similar to Toulmin's claim, CER evidence is similar to Toulmin's data, and CER reasoning is similar to Toulmin's warrant and backing (Toulmin, 2003; McNeill *et al.*, 2006; Osborne and Patterson, 2011). The key aspect of making a claim is an assertion or statement to solve a problem or question (McNeill *et al.*, 2006; McNeill and Krajcik, 2012; Van Eemeren *et al.*, 2014). A claim is always in doubt and therefore cannot stand alone. Thus, the claim must be advanced by known data (Osborne and Patterson, 2011). This results in the second element of evidence which is based on scientific data that support the claim and that the scientific data have to be appropriate and sufficient (McNeill *et al.*, 2006; McNeill and Krajcik, 2012). Moreover, a claim can be supported with multiple pieces of evidence since complex problems often require multivariate argumentation (McNeill and Krajcik, 2012). The third element is reasoning which acts as a bridge between claim and evidence. Reasoning answers the question as to why pieces of evidence support a claim, which requires the use of scientific principles (McNeill *et al.*, 2006). All elements are equally important as an argument is only complete when all three elements are present, *i.e.*, when a claim is strengthened through evidence and justified by reasoning (McNeill and Krajcik, 2012).

Scaffolding

Scaffolding is widely used in science education (Belland, 2017; Yuriev *et al.*, 2017; Fan *et al.*, 2020; Luo *et al.*, 2020; Flynn, 2021). It refers to multiple ways of assistance that support students' learning and reasoning including modelling or targeted prompting (Van de Pol *et al.*, 2010; Kang *et al.*, 2014). Scaffolding has been defined by Wood *et al.* as a process that "enables a child or a novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (Wood *et al.*, 1976, p. 90). Vygotsky (1980) built upon this idea when conceptualising the zone of proximal development, which describes that learners should, ideally, work in the zone of proximal development in which they are not able to solve a task on their own but with assisted guidance. A scaffold in this sense serves as an enabler for learners to complete tasks that they cannot solve alone (Wood *et al.*, 1976; Pea, 2004; Lajoie, 2005; Kang *et al.*, 2014). However, if a scaffold provides too much information, or structures the problem-solving process to a very large extent, it may no longer be as effective as students do not feel challenged (McNeill *et al.*, 2006). As a result, the support provided is strongly dependent on the individual (Van de Pol *et al.*, 2010). This arises the need for adaptive scaffolding as we acknowledge that the students' prior knowledge and former experience often differ significantly (Vygotsky, 1980; Pea, 2004). This is supported by a scaffold's core criterion of contingency including a tailored and differentiated support (Van de Pol *et al.*, 2010). Thus, an adaptive scaffold can be supportive for students' self-regulated learning since the students' individual needs are integrated in the design of the scaffold (Azevedo *et al.*, 2004).

As scaffolding is a temporary process assisting students if they need support, it is also important to fade out the given support when it is no longer needed (Lajoie, 2005; McNeill *et al.*, 2006). However, appropriate fading of support depends on the tasks' complexity and students' progress (Kang *et al.*, 2014). Fading too early may have adverse effects as students might not have yet fully understood certain concepts or activities (Noroozi *et al.*, 2017).

Scaffolding argumentation

Scaffolding can take place in a wide variety of disciplines and is also finding increasing use as a support for argumentation. In a meta-analysis of argumentation in computer-supported collaborative learning, Wecker and Fischer (2014) revealed that the occurrence of specific functional components (*i.e.*, whether and how often claim-evidence-reasoning occur in learning material) has a significant medium to large effect (Cohen's d = 0.72) on argumentation (Wecker and Fischer, 2014).

In addition to the benefits of using argument components, building strong arguments also involves the appropriate use of scientific concepts (Sandoval and Millwood, 2005; Choi *et al.*, 2013; Lieber and Graulich, 2022). Therefore, CER scaffolding does not only provide support for the structure of arguments but can also be enhanced with the incorporation of concepts (McNeill *et al.*, 2006; Songer and Gotwals, 2012). The interplay of concept knowledge and argumentation was demonstrated by Songer and Gotwals (2012) as students' conceptual understanding increased by using CER scaffolding in a pre-post intervention. In addition to the main components of CER scaffolds such as argumentation and concept knowledge, the type of scaffold should also be considered.

Kang *et al.* (2014) suggested that four of the six types of instructional scaffolding (originally analysed for English language learners (Walqui, 2006)) fulfil different functions in the construction of evidence-based explanations in a scientific context. These different types of scaffolds can be combined to be beneficial for the students. The first type is *instructional modelling*, which gives students clear examples of what is expected from them. This is especially important when implementing a new principle or task. Accordingly, students need to see in advance what the finished product should look like (Walqui, 2006; Kang *et al.*, 2014). In terms of argumentation and concept knowledge, this can be achieved by providing students with examples that illustrate the structure of an argument or presenting arguments that connect concept knowledge. The second type is *bridging* which functions as a link between existing and new knowledge (Walqui, 2006). This can be accomplished in scaffolds by asking students targeted questions that both activate their prior knowledge and link it directly to new content. The third type is *contextualising,* and this refers to using language in appropriate context as academic language is not only different from everyday language but often also intangible. For example, pictures or films can be used to support contextualisation (Walqui, 2006). When implementing this type of scaffold in argumentation, the visualisation of problem contexts or the illustration of the argument structure (in partial steps) is beneficial. The last type of scaffolding is *developing metacognition*, which fosters students' ability to evaluate and reflect on their current state of knowledge (Walqui, 2006). This function of a scaffold can be realised by prompting students to assess the difficulty and confidence, but also evaluating tasks and whether the students need help in certain areas of the task.

Research questions

Since building well-grounded arguments with the use of appropriate scientific principles is an essential skill for organic chemistry, our goal was to design an adaptive scaffold that supports students in argumentation based on their needs, *i.e.*, argumentation and/or concept knowledge support. Therefore, students were assigned to one of four groups of an adapted scaffold based on their performance on argumentation and concept knowledge in a diagnostic scaffold. This resulted in the overarching question whether the adaptive scaffold addresses students' experience of difficulties and closes the performance gap in a pre-post comparison. This overarching question was divided into two main research questions.

(1) To what extent does the adaptive scaffold close the gap in students' performance?

- (a) To what extent do the group performances differ after the diagnostic scaffold based on scoring argumentation and concept knowledge?
- (b) Do the group performances converge after the adapted scaffolds?

(2) Does the adaptive scaffold improve students' performance in the respective area of support (argumentation and/or concept knowledge)?

To answer these two research questions, a quantitative analysis based on students' answers of the diagnostic scaffold and adapted scaffolds was conducted.

Methods

Participants and study setting

The study was conducted at a private, liberal arts, research-intensive university in the North-eastern United States in April and May 2021. Sixty-four students were recruited voluntarily from the Organic Chemistry II course *via* an announcement in the lecture. Students received extra credit for their participation in the study, which was 1% of their final course grade. Organic Chemistry I is a prerequisite course for taking Organic Chemistry II. The Organic Chemistry I and II courses cover topics such as structure and properties of molecules, nucleophilic substitutions, carboxylic acids and their derivates, conjugated π-systems, retrosynthetic analyses and synthetic strategies. It is assumed that students are familiar and knowledgeable with these organic chemistry topics as they were discussed in class before the study was conducted. The students enrolled in the course were between 18–22 years of age and were majoring in chemistry, biochemistry, chemical engineering, and biology, among others. Among the students that participated in this study, 34, 29, and 1 student identified themselves as female, male, and non-binary, respectively, of which 56.3%, 31.3%, 4.7%, 4.7%, 1.6%, and 1.6% identified as white, Asian, Black, more than one race, Latino/a/x, and other races, respectively. All students created a user code as a pseudonym. The pseudonyms assigned to participants for the study as well as given names in this manuscript do not reflect their race, ethnicity, gender, or other identities.

Research instrument

To investigate how adaptively scaffolding students' argumentation patterns and concept knowledge affects students' argumentation skills in organic chemistry, we designed a sequence of two online scaffolds using Qualtrics, with each scaffold consisting of multiple parts (see ABC in Fig. 1). Students had 48 hours to complete each of the scaffolds with a completion time of approximately 60 minutes. The goal of the diagnostic scaffold was to establish a baseline of students' argumentation skills and their use of concept knowledge. Based on how the students performed in the diagnostic scaffold, they received one out of four versions of an adapted scaffold that focuses on supporting them in the area in which they had experienced the greatest difficulties in the diagnostic scaffold. Based on our previous study (Lieber and Graulich, 2022), we assumed that the students either experience challenges with (1) the appropriate use of argument components (claim, evidence, reasoning), (2) the appropriate use of chemical concepts in their argumentation, (3) with both, the appropriate use of argumentation components and chemical concepts, or (4) the ability to build multivariate arguments. The tasks underlying the scaffolds followed our previously published task design (Lieber and Graulich, 2020; Lieber and Graulich, 2022), in which students are prompted to judge the plausibility of alternative reaction pathways, which are shown in Fig. 2. Exemplary arguments for the eight alternative reaction pathways are shown in Appendix 1.

Diagnostic scaffold. The diagnostic scaffold used was the same for all students. It served as a pre-measure to determine the differences and difficulties that students have with argumentation on the plausibility of alternative organic reaction pathways. First, students were asked to predict the product of the reaction of 4-chlorobutanol with hydroxide and to justify their decision (see Fig. 2). This was followed by a task on general argumentation patterns. Here, students received the input that an argument consists of three basic components (claim, evidence, and reasoning), what function each component has, and a concrete example of how to use them in the building of arguments (see Step A, Fig. 1). It was also explicitly discussed that a claim can be supported with several pieces of evidence and that a piece of evidence can be justified with several reasoning statements. Following this explanation, students completed three tasks in which they were given a science-related argument and had to assign the basic components (claim, evidence, reasoning) to the appropriate sentence component. Students always had the option to indicate that they do not know which basic component can be assigned to which sentence component (see Appendix 2). After the task, students received sample solutions with accompanying explanations. The next step was the activation of chemical concepts related to the reaction of 4-chlorobutanol with hydroxide, where

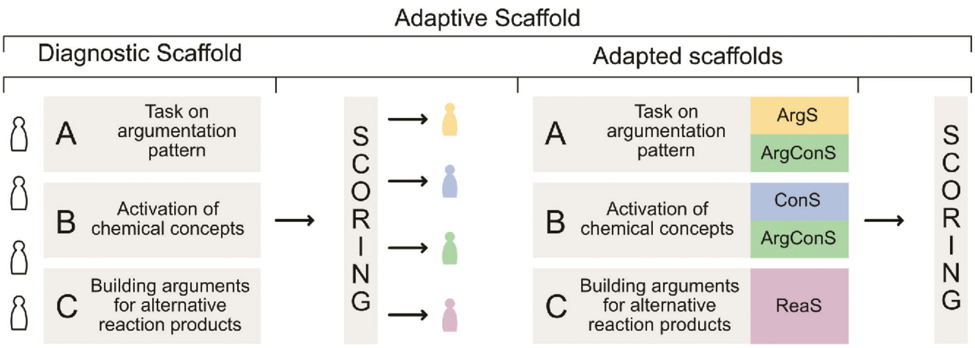


Fig. 1 Illustration of the areas of support for the four adapted scaffold groups. Dependent on their scoring after the diagnostic scaffold, the groups received an adapted scaffold for argumentation (ArgS group in yellow), use of concept knowledge (ConS group in blue), argumentation and use of concept knowledge (ArgConS in green), or multivariate reasoning (ReaS group in purple).

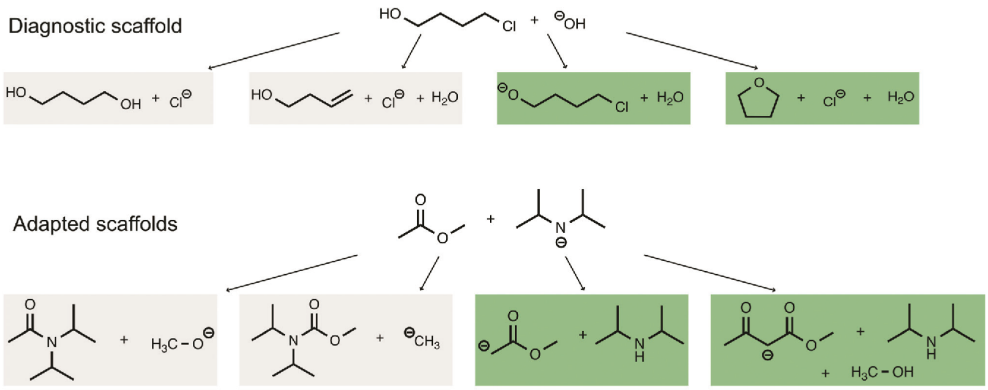


Fig. 2 Students judged the plausibility of four alternative reaction products for the reaction of 4-chlorobutanol and hydroxide in the diagnostic scaffold and for the reaction of methyl acetate with diisopropylamide in the adapted scaffolds. The correct reaction products are highlighted in green. There are two correct products for each reaction because the left green molecules are the precursors of the correct products.

students were given directed questions covering different general chemical concepts, *e.g.*, to decide whether steric aspects need to be considered in the reaction or to determine at which positions the molecules involved react as a nucleophile/an electrophile/an acid/a base (see Step B, Fig. 1 and Appendix 3). The questions were chosen to apply to any type of alternative reaction pathway students will encounter. After the argumentation pattern task and the activation of concept knowledge, students judged the plausibility of four alternative reaction products for the reaction of 4-chlorobutanol and hydroxide (see Step C, Fig. 1 and Appendix 4). In the first step, they stated their claim by deciding whether they think the reaction product shown is plausible or implausible. In a second step, they had to support the given claim with evidence and to justify it with reasoning. After building arguments for four alternative reaction products, the students had the opportunity to revise or defend their initial claim by indicating whether they still choose the reaction product they formed at the beginning of the diagnostic scaffold or whether they choose one or more of the alternative reaction products. The diagnostic scaffold ended with evaluation questions related to the argumentation task and the questions on chemical concepts.

Adapted scaffolds. The adapted scaffolds are comparable in style to the diagnostic scaffold. All four adapted scaffolds began with students predicting the product of the reaction of methyl acetate with diisopropylamide and justifying their decision (see Fig. 2). To repeat the structure of an argument, students received an argumentation task with new science-related arguments. Students were then given the same questions on chemical concepts as in the diagnostic scaffold, but answered them on the reaction of methyl acetate with diisopropylamide. They were also given four alternative reaction products for which they formed a claim, evidence, and reasoning. Lastly, students indicated whether they wanted to keep or revise their formed reaction product.

As described shortly above, the four adapted scaffold groups differed with regard to the additional support provided to the respective group of students.

Group 1. Argumentation support (ArgS). The ArgS group received additional support for structuring and building arguments, *i.e.*, the three components claim, evidence, and reasoning. The students were given eight different science-related arguments, divided into sentence sequences, for which they had to decide whether it was a claim, evidence or reasoning. Moreover, by building arguments by themselves when judging the plausibility of alternative reaction products, a definition of the argument components (claim, evidence, and reasoning) was added at the positions when their formation was required.

Group 2. Concept knowledge support (ConS). The ConS group obtained additional support in using appropriate chemical concepts when building arguments. The targeted questions on chemical concepts prior to the building of arguments remained unchanged, as students were first asked to activate their prior knowledge to link it to the reaction in the task. When building their arguments for the alternative reaction products, students received author-generated answers to the previous concept questions but had to interpret them by themselves, *e.g.*, they received the pK_a values of the molecules but no information about the interpretation of pK_a values.

Group 3. Argumentation + concept knowledge support (ArgConS). The ArgConS group received both, additional support for the structure of arguments from the ArgS group and additional support for the use of chemical concepts from the ConS group.

Group 4. Multivariate reasoning support (ReaS). Since the ReaS group already had built appropriate arguments including given if no chemical concepts were used. One point was awarded if at least 50% of the concepts were incorrect and two points were awarded if more than 50% of concepts were used correctly. In total, the concept knowledge score consisted of 29 points.

Step II: grouping into the four adapted scaffold groups. To divide the students into one of the four adapted scaffolding groups (support for argumentation (ArgS), support for concept knowledge (ConS), support for argumentation and concept knowledge (ArgConS), and support for multivariate reasoning (ReaS)), the argumentation score and the concept knowledge score were determined for each student. Here, a qualitative discussion within the research team took place for each student. We chose to use a percentage of the total score as a threshold. The percentage is 65% of the total score as it corresponds to the passing D grade in many US universities. Students who did not reach the threshold score received further support in argumentation patterns and/or the use of concept knowledge in argumentation. The threshold for the argumentation and concept knowledge score corresponds to 16 points (of 25 points) for the argumentation score and 18 points (of 29 points) for the concept knowledge score. These scores were continuously assessed that each student received the support that corresponded to their performance. Fig. 3 illustrates the grouping of the four adapted scaffold groups. Students were assigned to the ArgS group when the argumentation score was less than 16 and the concept knowledge score was at least 18. The ConS group included students whose argumentation score was at least 16 and the concept knowledge score was less than 18. When the argumentation score was less than 16 and the concept knowledge score was less than 18 a student was assigned to the ArgConS group. Students whose scores exceeded the threshold of 16 (argumentation score) and 18 (concept knowledge score) were dedicated to the ReaS group.

Step III: quantitative analysis of students' scores of the diagnostic scaffold and adapted scaffolds. Statistical measurement methods were used to answer the research questions and were conducted using the software R. Therefore, students' answers of the adapted scaffolds were scored in the same way as the answers of the diagnostic scaffold. The numbers and types of arguments in the task on argumentation patterns (see Step A, Fig. 1) differed in the adapted scaffolds. Therefore, we decided to only use students' arguments built for the alternative reaction pathways (see Step C, Fig. 1 and Appendix 4) to determine the argumentation score. Thus, the argumentation score of the diagnostic scaffold was also adapted by only using students' answers of Step C to be able to compare the argumentation score of the diagnostic scaffold and the adapted scaffolds. As a result, the new maximum argumentation score for the diagnostic scaffold and the adapted scaffolds consisted of 16 points, whereas the maximum concept knowledge score remained 29 points.

First, a visual inspection of the normal distribution of students' argumentation and concept knowledge scores was performed, which was supported by a Shapiro–Wilk test. This analysis revealed that the data were not normally distributed, which indicated the use of non-parametric tests. For all measurements, an α-level of 0.05 was used.

To determine to what extent the group performances differ after the diagnostic scaffold and whether the group performances converge after the adapted scaffolds, a Kruskal–Wallis test and subsequent *post hoc* comparisons with Wilcoxon rank-sum tests and Bonferroni-adjusted p-values were performed (Field *et al.*, 2012). The Kruskal–Wallis test as the non-parametric counterpart of the one-way ANOVA was chosen to compare more than two independent samples of different sample sizes. As the Kruskal–Wallis test only indicates whether groups are significantly different, subsequent *post hoc* comparisons are necessary to identify these groups. In case of

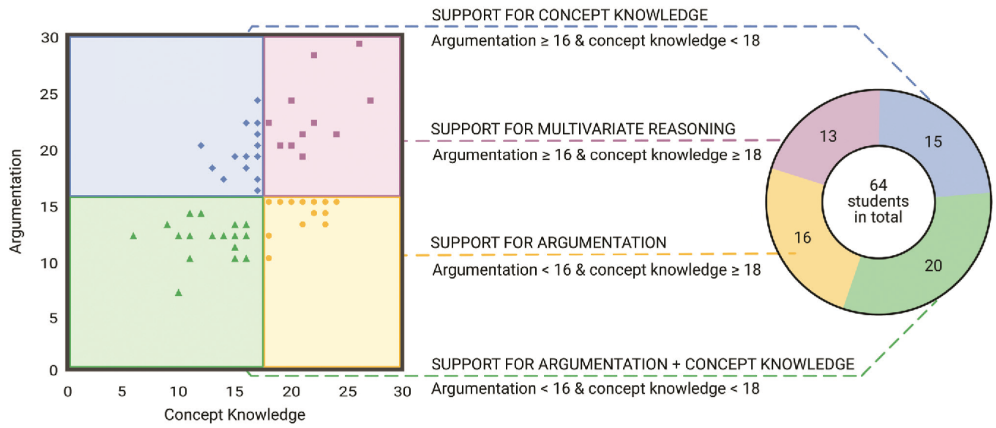


Fig. 3 Grouping of the four adapted scaffold groups.

significant results in the *post hoc* comparisons, the correlation coefficient r as a measure of effect size was calculated from the conversion of the z-score (Rosenthal, 1991). The correlation coefficient r was defined as $0.10 \leq r \leq 0.30$ as small effect, $0.30 \leq r \leq 0.50$ as medium, and $r \geq 0.50$ as large (Cohen, 1992).

To determine whether the adaptive scaffold improved students' performance in the areas of support, a Wilcoxon signed-rank test with Bonferroni-adjusted p-values was performed. The Wilcoxon signed-rank test as the non-parametric counterpart of the dependent t-test was chosen since we wanted to compare two dependent samples, *i.e.*, pre-post comparisons of the same group. In the case of significant results, the correlation coefficient r was reported as the effect size.

Results and discussion

RQ1: To what extent does the adaptive scaffold close the gap in students' performance?

To determine the extent to which an adaptive scaffold closes the gap in students' performance, a Kruskal–Wallis test was performed. The analysis showed whether the groups significantly differ after the diagnostic scaffold, which serves as a pre-measure, to verify the qualitative grouping of the students. In case of significant results, subsequent *post hoc* comparisons with Wilcoxon rank-sum tests and Bonferroni-adjusted p-values were performed. The statistical results are summarised in Table 1.

After the diagnostic scaffold (pre-measure), the groups differed significantly in terms of their argumentation score $H(3) = 43.97$, $p = <0.001$. *Post hoc* comparisons revealed significant differences with large effects in four of six group comparisons, indicated with black lines on the left side in Fig. 4. However, the ArgS group (yellow) and the ArgConS group (green) as well as the ConS group (blue) and the ReaS group (purple) did not vary significantly in the pre-measure. This sheds light on the appropriateness of the qualitative grouping since in the pre-measure, both groups (ArgS and ArgConS) with students' argumentation scores below the threshold of 16, differed significantly with a large effect from the two groups (ConS and ReaS) in which students' argumentation score was above the threshold and, thus, from those groups who will not receive additional argumentation support. Moreover, the pre-measure of the ArgS and ArgConS groups, who will receive an adapted

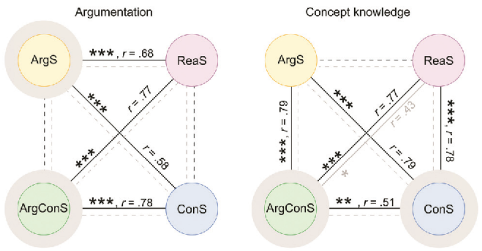


Fig. 4 Differences between the four groups after the diagnostic scaffold (shown in black lines) and after the adapted scaffold (shown in grey lines). In case of non-significant differences, the black and grey lines are dashed. Groups who received an adapted scaffold in the respective area are highlighted with a grey background circle. Significance levels of the group comparisons are indicated (*p < 0.05; **p < 0.01; ***p < 0.001).

scaffold for argumentation patterns, did not vary significantly from each other. The ConS and ReaS groups also did not differ significantly after the diagnostic scaffold and will consequently not receive an adapted scaffold for argumentation patterns. Fig. 4 illustrates the non-significant comparisons after the diagnostic scaffold (pre-measure) with dashed black lines. These results indicate that the qualitative grouping was successful for the argumentation score because the ArgS and ArgConS groups, who will receive an adapted scaffold for argumentation patterns, differed significantly from the two groups, ConS and ReaS groups, not receiving additional support for argumentation patterns.

After the diagnostic scaffold (pre-measure), the groups were found to differ significantly in terms of their concept knowledge score $H(3) = 50.92$, $p = <0.001$. Subsequent *post hoc* comparisons were performed and demonstrated that five of the six group comparisons differed significantly with large effects, which is illustrated with black lines on the right side of Fig. 4 and summarised in Table 1. Comparison of the concept knowledge score revealed that the two groups, ArgS and ReaS, who will subsequently not receive additional information on chemical concepts, did not differ significantly. The two groups (ConS and ArgConS), who will receive an adapted scaffold on concept knowledge, varied from each other in the pre-measure. Thereby, the ConS and ArgConS groups are significantly different with a large effect. The fact that both

Table 1 Results of the *post hoc* comparisons with Bonferroni-adjusted p-values of the Kruskal–Wallis test of the argumentation score and concept knowledge score after the diagnostic scaffold. The correlation coefficient r is reported as effect size in case of significant p-values. Significant results are highlighted in bold

	Diagnostic scaffold argumentation score				Diagnostic scaffold concept knowledge score			
Comparisons	$M_{\text{first group}}$	$M_{\text{second group}}$	p	r	$M_{\text{first group}}$	$M_{\text{second group}}$	p	r
ArgS *vs.* ConS	10	13	**0.001**	0.58	21	16	**<0.001**	0.79
ArgS *vs.* ArgConS	10	8.5	0.056		21	13.5	**<0.001**	0.79
ArgS *vs.* ReaS	10	14	**<0.001**	0.68	21	21	>0.999	
ConS *vs.* ArgConS	13	8.5	**<0.001**	0.78	16	13.5	**0.003**	0.51
ConS *vs.* ReaS	13	14	>0.999		16	21	**<0.001**	0.78
ArgConS *vs.* ReaS	8.5	14	**<0.001**	0.77	13.5	21	**<0.001**	0.77

groups, who will receive an adapted scaffold on concept knowledge (ConS and ArgConS), differ from each other, but also that these groups differ from the two other groups (ArgS and ReaS), who will not receive an adapted scaffold on concept knowledge, revealed that the ConS and ArgConS groups received their adapted scaffold on a legitimate basis. This means that the qualitative grouping was also successful for the concept knowledge score. However, it is not surprising that the ArgConS group is significantly different from all of the other groups since the students received distinctly fewer points in the concept knowledge score. These differences in score can also be observed through qualitative observations of the students' answers, as many questions (*e.g.*, on nucleophilicity and electrophilicity, steric effects, or electronic effects) were either answered incorrectly or with the phrase "I don't know."

After analysing the differences in students' performance after the diagnostic scaffold (pre-measure), the results of the Kruskal–Wallis test and subsequent *post hoc* comparisons of the group performances after the adapted scaffolds (post-measure) were reported. In terms of the argumentation score, the groups did not differ significantly in the post-measure $H(3) = 4.79$, $p = 0.188$. This is illustrated with dashed grey lines in Fig. 4. Therefore, it can be assumed that after the adapted scaffolds, the groups have converged in terms of their argumentation score, which means that no significant differences were measurable. Regarding the concept knowledge score, the groups differed significantly from each other in the Kruskal–Wallis test $H(3) = 10.46$, $p = 0.015$, but in the subsequent *post hoc* tests with Bonferroni-adjusted p-values it became clear that only two groups differ significantly from each other after the adapted scaffolds (post-measure). The ArgConS and ReaS groups still differ from each other significantly with a medium effect after the adapted scaffolds in the concept knowledge score, which is illustrated in Fig. 4 with a grey line. All other group comparisons did not show a significant difference in *post hoc* comparisons, shown in Table 2. This reveals that the four groups also converged in terms of the use of concept knowledge and that the performance gap which was diagnosed beforehand is not significantly noticeable in the post-measure. This suggests overall that the adaptive scaffold supported the respective groups and closed the gap in their performance. The difference between the ArgConS and ReaS groups can be understood when considering them as two sides of a continuum. Students in the ArgConS group received the lowest scores in the concept knowledge score whereas the students of the ReaS group achieved the highest scores in this category. This is also noticeable in the pre-measure since the ConS and ArgConS groups differed significantly although both groups will receive an adapted scaffold for concept knowledge. However, these two extrema came a little closer to each other demonstrated by the comparison between the median values from the diagnostic scaffold (difference of median values for the ReaS group and ArgConS group = 7.5) and the adapted scaffolds (difference of median values for the ReaS group and ArgConS group = 7).

Table 2 Results of the *post hoc* comparisons with Bonferroni-adjusted p-values of the Kruskal–Wallis test of the concept knowledge score after the adapted scaffolds (post-measure). The correlation coefficient r was reported as effect size in case of significant p-values. Significant results are highlighted in bold

	Adapted scaffolds concept knowledge score			
Comparisons	$M_{\text{first group}}$	$M_{\text{second group}}$	p	r
ArgS *vs.* ConS	20	19	>0.999	
ArgS *vs.* ArgConS	20	18	>0.999	
ArgS *vs.* ReaS	20	25	0.543	
ConS *vs.* ArgConS	19	18	>0.999	
ConS *vs.* ReaS	19	25	0.188	
ArgConS *vs.* ReaS	18	25	**0.013**	0.43

Overall, it became apparent that the groups who will receive additional support in the adapted scaffolds were significantly different to those who will not receive support, with a large effect in the pre-measure. Thus, this confirms that the grouping after the diagnostic scaffold was successful and revealed that a performance gap was present. After the adapted scaffolds (post-measure), the groups did not differ significantly in terms of argumentation score and concept knowledge score. The ArgConS group constitutes an exception as this group differed significantly from all the other groups regarding the concept knowledge score in the pre-measure and the ReaS group in terms of the concept knowledge score in the post-measure. This is because the ArgConS group was conceptually weaker after both the diagnostic scaffold and adapted scaffold compared to all three other groups.

RQ2: Does an adaptive scaffold improve students' performance in the respective area of support (argumentation and/or concept knowledge)?

To compare the performance of each group in terms of argumentation score and concept knowledge score, a Wilcoxon signed-rank test with Bonferroni-adjusted p-values was performed. Fig. 5 summarises and compares the performance of each group in the diagnostic scaffold (pre) and adapted scaffolds (post). Detailed results are shown in Appendix 5. The following examples represent students' original built arguments. Thereby, all students received free-text boxes labeled as evidence and reasoning, which allows students to independently decide the assignment of statements to evidence and reasoning, as well as the numbering and order of the statements (see Appendix 4, Fig. 8). Therefore, certain pieces of evidence and reasoning statements do not fulfill all required characteristics of argument components. Moreover, certain statements are not technical correct regarding scientific principles. However, these students have never built arguments before, hence, it is not surprising that they did not build perfect arguments. Instead, they improved their argumentation during the adaptive scaffold.

The ArgS group increased significantly in the argumentation score from pre- ($M = 10$) to post- ($M = 14.5$) measure ($V = 107.5$, $p = 0.007$, $r = 0.68$) with a large effect. For the concept knowledge score, on the other hand, there was no significant change from pre ($M = 21$) to post ($M = 20$) ($V = 47.5$, $p = 0.299$)

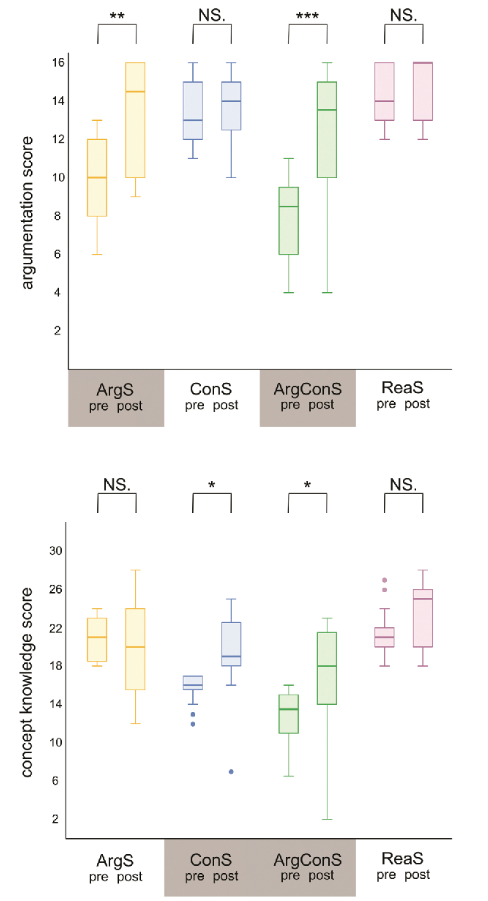


Fig. 5 Argumentation scores and concept knowledge scores of the four groups after the diagnostic scaffold (pre) and adapted scaffolds (post). Groups who received an adapted scaffold in the respective area are highlighted with a grey background. Horizontal stripes in box plots indicate median-values. Significance levels of the pre- and post-comparisons are indicated (NS. $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

measurable. Therefore, the adapted scaffold on argumentation patterns supported the students significantly in building arguments but not in the use of concept knowledge. This result is encouraging as the students only received additional support on argumentation patterns. Louis, a participant in this study, serves as a student example of the ArgS group. He received 15/25 points in the argumentation score and 24/29 points in the concept knowledge score, which resulted in assigning him to the ArgS group. After receiving an adapted scaffold for building arguments, Louis received 16/16 points in the argumentation score and 27/29 in the concept knowledge score. In the diagnostic scaffold, he claimed that tetrahydrofuran (THF) is an implausible product of the reaction from 4-chlorobutanol and hydroxide. For that purpose, Louis built his argument by using the free-text boxes labeled as evidence and reasoning as follows:

Claim: *The reaction product is implausible.*

Evidence 1: *While it is plausible that the OH will initially deprotonate the alcohol, the molecule must contort significantly to put the alkoxide group next to the C–Cl bond.*

Reasoning 1: *The likelihood of getting the alkoxide group next to the C–Cl bond of the same molecule before, say, the alkoxide reacts with the C–Cl bond of a neighboring identical molecule is unlikely because the molecule would have to bend in on itself. In addition, entropics don't favor the reaction because a ring has less entropy than a straight chain.*

It is noticeable that Louis, referring to the argument structure, related the evidence to the claim and even explicitly mentioned what he claimed implausible regarding the formation of THF. However, after that, he did not clearly separate the argument components evidence and reasoning. In the first part of his reasoning statement, he described the evidence again in other words, but did not explain why "the molecule must contort significantly". Moreover, in the second part of his reasoning statement, there is another piece of evidence and corresponding reasoning statement regarding entropy.

In the adapted scaffold, Louis then received additional argumentation support and built the argument shown below for the reaction of methyl acetate and diisopropylamide (LDA) to methyl acetoacetate *via* a Claisen condensation.

Claim: *The reaction product is plausible.*

Evidence 1: *The amine is more likely to act as a base than nucleophile.*

Reasoning 1: *The amine is sterically bulky so it can't approach an electrophile easily but can deprotonate a molecule.*

Evidence 2: *The amine is basic.*

Reasoning 2: *Since the amine has a negative charge and doesn't have resonance structures it will be stabilized by receiving a proton and become neutral.*

Evidence 3: *The ester has acidic alpha protons.*

Reasoning 3: *These protons are acidic because the enolate conjugate can be resonance stabilized.*

Evidence 4: *The enolate can attack an ester because the carbanion is nucleophilic and the carbonyl is electrophilic.*

Reasoning 4: *The carbanion is nucleophilic because it is negatively charged and the carbonyl is electrophilic because the oxygen pulls electron density away.*

At a first glance, it becomes clear that Louis has made a separation between evidence and reasoning. Each of his built pieces of evidence refers to the claim and answers the why-question. Furthermore, his reasoning statements consist of scientific principles that justify the evidence. By comparing these two formed arguments, Louis has improved significantly in building arguments by not only making a distinction between evidence and reasoning statements but also by building more pieces of evidence and reasoning, which gives depth to his argument. However, Louis can continue to improve in building arguments in the future. For example, he can concretize evidence 1 by already mentioning the structure of the amine. Furthermore, he can split evidence 4 and consider the nucleophilicity and electrophilicity separately from each other to address the electronic properties of the molecules more specifically in his reasoning statements.

When looking at the ConS group, the data analysis revealed a comparable trend, as in this group, there was no significant change for the argumentation score (pre-M = 13, post-M = 14), V = 30, p = 0.823. In comparison the concept knowledge score increased significantly from pre (M = 16) to post (M = 19), with a large effect (V = 82, p = 0.011, r = 0.66). This means that the adapted scaffold on the use for concept knowledge supported the students significantly for using concept knowledge but not in building argumentation patterns. This result is also encouraging as students in the ConS group only received additional support in the use of concept knowledge. Jessica, a participant in this study, serves as an example for the ConS group since she received 21/25 points in the argumentation score but 17/29 points in the concept knowledge score. After she worked with an adapted scaffold for the use of concept knowledge, she obtained 13/16 points in the argumentation score and 25/29 points in the concept knowledge score.

Jessica's example argument is also regarding the formation of THF as a reaction product from 4-chlorobutanol and hydroxide.

Claim: *The reaction product is plausible.*
Evidence 1: *It forms stable products.*
Reasoning 1: *The Cl is more stable than the OH, H_2O is more stable than OH, and a membered ring with oxygen is relatively stable.*
Evidence 2: *There could be conditions where this is the most favorable reaction.*
Reasoning 2: *Though the diol attacking itself to form a ring is likely not the most kinetically favorable reaction, sometimes certain conditions promote reactions like that.*

Jessica supported her claim with pieces of evidence by stating that the products are stable and that the reaction would be possible under certain conditions. Both pieces of evidence are rather vague and do not give a concrete indication for which conceptual reasons the formation of THF is plausible. While Jessica has not used any incorrect concepts *per se*, she did not elaborate on them. Instead, she justified stable products by comparing stability, but without elaborating on the reasons for stability. She also did not specify the reaction conditions. However, after getting additional information on chemical concepts in the adapted scaffold, such as the pK_a values of the involved molecules or the electronegativity values, she built the following argument on the formation of methyl acetoacetate from methyl acetate and LDA.

Claim: *The reaction product is plausible.*
Evidence 1: *The product is not entropically unfavorable.*
Reasoning 1: *3 molecules become 3 molecules*
Evidence 2: *The negative charge on the product is stabilized by resonance.*
Reasoning 2: *It can put negative charge on two different oxygens (allylic to two C=O bonds).*
Evidence 3: *The negative charge is stabilized by inductive effects.*
Reasoning 3: *It is allylic to two C=O bonds, which are electron-withdrawing.*
Evidence 4: *The ester will be more likely to donate a proton.*
Reasoning 4: *It has a lower pK_a than the amine.*
Evidence 5: *The final products are all stable.*
Reasoning 5: *Methanol and the amine are both stable as well as the ester and enolate compound.*

Jessica, like Louis in the ArgS group, improved the quality of her argument with respect to the concept knowledge used. With the help of the additional concept information, which she still had to interpret herself, Jessica used concepts such as entropy, electronic effects, and acidity. In contrast to the diagnostic scaffold, she did not remain vague but used explicit scientific principles to support and justify her claim. Jessica tried to include a variety of scientific principles in building her arguments, since, for example, the consideration of entropy alone would not have been a sufficient justification. Furthermore, she did not build reasoning statements by repeating comparable statements she already used in her pieces of evidence, which is still apparent in reasoning 5. This is an evident improvement in her performance after the adapted scaffold. Thus, Jessica could continue to improve in building arguments in the future. In evidence 5 and reasoning 5, it becomes apparent that she should engage again with the concept of stability, as she is unable to provide a satisfactory justification for the stability of molecules in both the diagnostic scaffold and the adapted scaffold.

The ArgConS group, who received support in both, argumentation and concept knowledge, was able to achieve a significant increase with a large effect in both, the argumentation score (pre-M = 8.5, post-M = 13.5, V = 193.5, p = <0.001, r = 0.74) as well as in the concept knowledge score from pre- to post-testing (pre-M = 13.5, post-M = 18, V = 158.5, p = 0.011, r = 0.57). The student Mike is exemplary of the ArgConS group; he received 10/25 points in the argumentation score and 16/29 points in the concept knowledge score in the diagnostic scaffold. After working with the adaptive scaffold on building arguments and using concept knowledge, he obtained 16/16 points in the argumentation score and 23/29 points in the concept knowledge score. In Mike's example argument, he claimed that an alkoxide is a plausible product of the reaction of 4-chlorobutanol and hydroxide.

Claim: *The reaction product is plausible.*
Evidence 1: *The oxygen is better stabilized.*
Reasoning 1: *On a larger molecule, the negative charge is better stabilized because it can be stabilized through resonance.*
Evidence 2: *The smaller molecule is more stable.*
Reasoning 2: *A water molecule is much more stable than a hydroxyl group.*

Like most of his fellow students, Mike tried to build a reasoning statement for each piece of evidence. Evidence 1 meets all conditions of an evidence statement, such as the answer to the why question and that it is an explanation rather than a description. Reasoning 1 also formally meets the criteria, but technical deficiencies appear, for example, the alkoxide cannot be stabilized by resonance, which is a common misconception (Carle and Flynn, 2020). In the second part of the argument, Mike used the same argument as Jessica with respect to stability, as he tried to justify 'the stability of molecules with their stability' instead of justifying the stability with chemical concepts (see evidence 2 and reasoning 2). In reasoning statement 2, he did not provide any further information, so the justification of the evidence is not apparent. Furthermore, he remained vague in evidence 2 because without the reasoning statement it is not clear which molecule he referred to as "smaller molecule". In the adapted scaffold, Mike worked on an adapted scaffold for building arguments *and* using concept knowledge, building the following argument for the formation of methyl acetoacetate from methyl acetate and LDA.

Claim: *The reaction product is plausible.*
Evidence 1: *The reaction is entropically favored.*
Reasoning 1: *There are more products than reactants which leads to increasing disorder.*
Evidence 2: *The oxygen is stable with the negative charge.*
Reasoning 2: *Oxygen is electronegative and can stabilize the negative charge after the molecule is rearranged via resonance.*
Evidence 3: *The nitrogen is not very stable with the negative charge.*
Reasoning 3: *The nitrogen is not very electronegative and thus cannot stabilize the charge as well.*
Evidence 4: *The negatively charged product is larger.*
Reasoning 4: *The negative charge can be better stabilized.*
Evidence 5: *The ester is a better leaving group.*
Reasoning 5: *The negatively charged oxygen is protonated, which makes its formation more favorable.*

From an argument structure point of view, Mike improved considerably. All of his built pieces of evidence supported the claim and all reasoning statements justified the evidence. He also used scientific principles as justification. From a conceptual point of view, Mike has improved, but this does not mean that all statements were technically correct. For example, he talked about the reaction being entropically favoured due to a higher number of products compared to reactants, which is incorrect because there are three molecules involved in the reaction on both the reactant and product side. Furthermore, in evidence 5, he referred to an ester as a leaving group. Here it is impossible to understand which molecule Mike was referring to as the leaving group since no ester is split off during the reaction. Nevertheless, Mike's improvement is evident across all arguments, both in terms of argument structure and the use of concept knowledge. In the future, Mike could further be supported in building arguments. Thereby, he can concretize his justification, for example by providing a counterpart in comparisons (*e.g.*, X is a better leaving group than Y or molecule X is larger than molecule Y).

The scores in the ReaS group suggest that students in this group did not achieve a significant improvement in either the argumentation score from pre (M = 14) to post (M = 16), V = 34, p = 0.534, or the concept knowledge score from pre (M = 21) to post (M = 25), V = 51, p = 0.118, opposite to the other groups. This is not surprising because the students already received high scores in the diagnostic scaffold. Thus, the scores did not differ significantly which is a good result. Rachel is a student representative of the ReaS group. She received 23/25 points in the argumentation score and 26/29 points in the concept knowledge score in the diagnostic scaffold. For the formation of THF as a product of the reaction of 4-chlorobutanol and hydroxide, Rachel has built the following argument.

Claim: *The reaction product is plausible.*
Evidence 1: *Intramolecular reactions are faster than intermolecular ones.*
Reasoning 1: *Rate is dependent on concentration of substrate. When the reactants are connected, there is essentially limitless substrate and thus this reaction can take place quite quickly.*
Evidence 2: *Hydroxide will deprotonate the hydroxyl.*
Reasoning 2: *The* pK_a *of hydroxyl is similar to that of hydroxide and they will thus exist in a proton transfer equilibrium. When the alkoxide ion is formed it will react with the nearby electrophile.*
Evidence 3: *This reaction is enthalpically favorable.*
Reasoning 3: *Five membered rings are stable because they have optimal bond angles for* sp^3 *hybridization. This means that they are lower in energy/more stable and thus this reaction will be exothermic.*

All arguments built by Rachel can be used as sample solutions for other students. She was able to separate her pieces of evidence from her reasoning statements as the pieces of evidence supported her claim and the reasoning statements justified the pieces of evidence. Moreover, Rachel used several chemical concepts such as enthalpy, kinetics, and basicity in her argumentation. After the adapted scaffold, Rachel obtained 16/16 points in the argumentation score and 28/29 points in the concept knowledge score. The students of the ReaS group were the only ones who were additionally prompted to build up to three reasoning statements for one piece of evidence. For the formation of an enolate as a product of the reaction of methyl acetate and LDA, Rachel built the following argument.

Claim: *The reaction product is plausible.*
Evidence 1: *The amide anion is highly basic.*
Reasoning 1.1: *High electron density on nitrogen.*
Reasoning 1.2: *Nitrogen adjacent to electron donating groups.*
Reasoning 1.3: *Bulky groups mean it won't act as a nucleophile.*
Evidence 2: *The alpha proton is slightly acidic.*
Reasoning 2.1: *The negative charge from removing a proton will be resonance stabilized.*
Reasoning 2.2: *The carbonyl contributes an inductive effect.*
Evidence 3: *The reaction is enthalpically favorable.*
Reasoning 3.1: *A weaker acid is formed than on the reactant side (protonated amide).*
Reasoning 3.2: *A weaker base is formed (enolate) than on the reactant side.*
Reasoning 3.3: *The enolate is resonance stabilized.*

Rachel improved in both, her argumentation score and concept knowledge score, but already performed well in the diagnostic scaffold. It is noticeable that her arguments are at a high level. Her evidence and reasoning statements answered the why-questions, and her use of scientific principles is multivariate. By building up to three reasoning statements, Rachel became more detailed and justified her pieces of evidence from multiple perspectives. In her first reasoning statement, for example, she justified the basicity of LDA given the electronic effects of nitrogen, and the electronic and steric effects of the adjacent groups. In Rachel's case, she could be further supported to include more scientific principles in her arguments in the future, *e.g.*, arguing with entropy or pK_a values.

In summary, the adapted scaffolds improved students' performance in the respective areas of support. The groups that received support in building arguments (ArgS and ArgConS) improved significantly with a large effect on the argumentation score and the groups. Those who received additional support in using concept knowledge (ConS and ArgConS) improved their performance significantly with a large effect on the concept knowledge score. Only the ReaS group showed no significant improvement, which is not surprising because the students of the ReaS group already had high scores in the diagnostic scaffold. Moreover, the fact that only the groups who received extra support in argumentation patterns and/or concept knowledge improved their performance in this area indicates that the improvement is not a simple training effect. Instead, the adaptive scaffold might be responsible for students' improvement.

Conclusions and implications

In organic chemistry, the use of chemical concepts and decision-making regarding alternative reaction pathways plays a central role in determining the reaction product. Building arguments following a structured method allows the appropriate use and reasoning of chemical concepts. However, there is a current gap between what students should learn and what they learn since the focus in chemistry is often on facts and memorisation (Stowe *et al.*, 2021), and building arguments is still rarely taught in their studies. The challenges faced by students with structuring arguments and appropriately using concept knowledge have been previously reported (Cruz-Ramirez de Arellano and Towns, 2014; Moon *et al.*, 2016; Deng and Flynn, 2021; Lieber and Graulich, 2022).

In this study, we investigated if adapted scaffolding that provides students with support in the area of argumentation and use of concept knowledge can make a significant difference in performance. Based on a diagnostic scaffold, which served as a pre-measure to analyse how students build arguments and how they used concept knowledge, students received an argumentation and concept knowledge score. They were then assigned to one of four adapted scaffolding groups: support for argumentation (ArgS), concept knowledge (ConS), argumentation and concept knowledge (ArgConS), and multivariate reasoning (ReaS). Consequently, each group received a different scaffold tailored to their needs in argumentation and/or use of concept knowledge. An argumentation score and concept knowledge score were given to each student based on their performance in the adapted scaffolds (post-measure).

The first research question in this study, examined to what extent the support groups differed from each other in the pre- and post-measure. When evaluating students' answers, it became apparent that the groups differed significantly from each other in the pre-measure. In particular, students grouped into the ReaS group already built well-grounded arguments after the diagnostic scaffold (pre-measure) as multiple pieces of evidence and reasoning were used as support and justification of the claim. Both evidence and reasoning statements consisted to a large extent of scientific principles and answered the why-questions. The other three groups still showed some gaps in the pre-measure in terms of argumentation and/or the use of concept knowledge. A closer look at the analysis revealed that the ArgS and the ArgConS groups each differed significantly with high effects (r = 0.58 and r = 0.78) in terms of the argumentation score from the ConS and the ReaS groups after the diagnostic scaffold. However, the ArgS and ArgConS group, as well as the ConS and the ReaS groups, did not vary significantly from each other, which exemplifies that the argumentation score determined group performance differences in building arguments but not in terms of concept knowledge. This indicated that the grouping of the students in the ArgS and the ArgConS groups in the adapted scaffolds for argumentation patterns was successful. By comparing the groups in terms of concept knowledge, the ConS and the ArgConS groups differed significantly from the ArgS and the ReaS groups with high effects (between r = 0.77 to r = 0.79) in the pre-measure. Therefore, the grouping in the adapted scaffolds for the use of concept knowledge was also successful. An exception is evident when comparing the ConS and ArgConS groups. Both groups differed significantly from each other with a high effect (r = 0.51) after the diagnostic scaffold, although both groups will receive an adapted scaffold for concept knowledge. However, this can be explained as students in the ArgConS group were conceptually weaker in comparison to all other groups. Based on this first analysis, the initial grouping was successful and the groups received the support in the respective area needed (*i.e.*, regarding argumentation and the use of concept knowledge). The second part of the first research question investigated whether the gap that occurred between the groups at the beginning could be closed using the adaptive scaffold. No significant differences between the groups were apparent in the argumentation skill after applying the adapted scaffold. In the area of concept knowledge, a significant difference with a medium effect (r = 0.43) only occurred between the ArgConS and ReaS groups. All other group comparisons showed no significant differences. It can therefore be assumed that the adaptive scaffolds closed the gap in students' performance. More chemical concepts were used to build arguments. The link between concept knowledge and the argument components is a key aspect in building arguments as it is considered an important part of the quality of an argument (Sandoval and Millwood, 2005; Choi *et al.*, 2013). In this context, one might assume that argumentation and concept knowledge can be fundamentally distinguished from each other. Songer and Gotwals (2012) reported a connection between concept knowledge and argumentation, which was not found in this study. However, one should not interpret this as evidence for the interdependence of argumentation and concept knowledge. The scoring process, separating argumentation and concept knowledge, did not explicitly acknowledge this linkage since strict attention was paid to the fact that both topics were considered separately from each other in the scoring process. Thus, in building arguments, technical correctness was not considered, and in the use of concept knowledge, no attention was given to whether the argument components (evidence and reasoning) were built correctly.

In the analysis of the second research question, the scoring results of the pre-post comparisons were compared to determine a possible improvement in each support group. Here, with respect to the argumentation score, the two groups that improved significantly with a high effect (r = 0.68 and r = 0.74) were those that received additional support for argumentation (ArgS and ArgConS group). Similarly, for the concept knowledge score, only the two groups that received support for the use of concept knowledge improved significantly with a high effect (r = 0.57 and r = 0.66) (ConS and ArgConS group). These results suggest that the adaptive scaffolds targeted areas where support was needed. Thus, this study demonstrates that an adaptive scaffold improved students' performance in the respective area of support. In organic chemistry, scientific reasoning on reaction pathways and products requires considering multiple chemical concepts in the decision-making process so that alternative reaction products and by-products can also be considered (Popova and Bretz, 2018). The implementation of this adaptive scaffold is useful in supporting students in applying the content to context, such as in suggesting alternative reaction products (Chen, 2014; Graulich and Caspari, 2021).

Implications for teaching

The results from this study suggest that the adaptive scaffold can be useful in teaching and learning organic chemistry. Thereby, teachers should consider adapting scaffolds to provide each student with the support they need. Teachers do not have to use the adaptive scaffold as a whole in their classes but can apply and adapt several parts. The scoring system does not necessarily need to be applied in class, which is the most time-consuming part of the scaffold. Thus, the possibility arises to separate tasks on argumentation patterns and the use of concept knowledge to support students first in building appropriate argument components (*i.e.*, claim, evidence, and reasoning). When students can build arguments, the formation can be enhanced by using chemical concepts appropriately. This might be beneficial as we did not observe an improvement in students' concept knowledge when students only received an adapted scaffold for argumentation pattern. Depending on students' prior knowledge regarding building arguments only the activation of concept knowledge can be used and adapted to the current course topic. Thereby, building arguments can also encourage students to strengthen and connect chemical concepts by applying it in building arguments. This indicates that the adaptive scaffold can be adapted to a variety of reactions and reaction mechanisms in organic chemistry. Moreover, by expanding the argumentation pattern by building several reasoning statements per evidence, students can be encouraged to use various chemical concepts in their argumentation. In addition, building arguments with this material can also be used collaboratively. Here, students can evaluate each other's arguments in peer-review processes. Thereby, students can undertake the task of scoring their peers. As a result, students experience a change in perspective that leads to an analytical decision-making process (Milkman *et al.*, 2009). Moreover, discussing the problem together can lead to an enhancing in understanding (Smith *et al.*, 2009) and the inhibition to make mistakes is reduced (Coppola and Pontrello, 2014). Additionally, students have the opportunity not only to build arguments against the product in the case of implausible reaction products but also to practice building counterarguments in the group discussions.

Implications for research

This study also has implications for research. First, in a further inquiry, interviews with students could be conducted to revise and improve the scaffolds. For example, to learn in more detail how students experienced the scaffolds, their thought processes while solving the problems, or the difficulties students encountered. In addition, the learning environment could be extended and applied over a longer period of time. Fading can then play a central role in this process. Here, it might be important to adaptively fade the scaffold according to students' learning progress, as the relevant content should be understood first (Kang *et al.*, 2014; Noroozi *et al.*, 2017). Fading also gives students greater responsibility for their learning over time (McNeill *et al.*, 2006). Moreover, when the scaffolds are conducted over a longer period of time, the effectiveness of the adaptive scaffold can be further examined. This might be beneficial for an improvement in students' concept knowledge as Songer and Gotwals (2012) revealed in their longitudinal study. Moreover, it calls us as researchers to purposefully rethink scaffolding because not all students benefit equally from a scaffold as they experience a variety of individual challenges (Caspari and Graulich, 2019; Petritis *et al.*, 2022). Thus, to provide students with even more individualised support, a computer-based adaptive learning system can be generated so that each student receives support exactly in the area needed. However, a large amount of data is required in advance (Zhou, 2016) and more adapted scaffolds need to be created to support content-related interpersonal differences (Shute and Zapata-Rivera, 2008) and acknowledge other individual differences such as motivational, metacognitive, and affective aspects, among others (Azevedo and Gasevic, 2019). An advantage of the computer-based adaptive learning system would be that the time and number of staff can be reduced and so more students have the opportunity to receive adaptive support in a shorter period of time (Dood *et al.*, 2020; Lee *et al.*, 2021; Yik *et al.*, 2021). A useful application would be an automatic scoring system of students' scaffold answers with machine learning approaches as scoring is the most time-consuming part of the adaptive scaffolding process.

Limitations

The conclusions discussed in the above paragraphs should be considered with caution as the study has certain limitations. First, the reactions used in the diagnostic and adapted scaffolds are only comparable to a certain extent because the organic chemistry reactions are different. However, it would not have been appropriate to use the same reaction twice as students may simply recall it in the adapted scaffolds. Both reactions were aligned well with the curriculum so that all required chemical concepts were covered in advance. Secondly, around one scoring point was used to assign the respective support group. However, apart from the score, each student was individually assessed as to the type of support needed based on the answers given. Lastly, we observed that students' answers were shorter in written arguments compared to oral interviews, which resulted in answers that were not always precise. To evaluate the students objectively and to guarantee impartiality, the students' answers were taken verbatim. An exception was when

View Article Online

students named, for example, molecules incorrectly, but one could make an unambiguous assignment based on the statement (*e.g.*, mistaking hydroxide with hydroxyl).

Author contributions

Leonie Sabine Lieber: conceptualisation, formal analysis, methodology, project administration, writing – original draft, visualization. **Krenare Ibraj**: methodology, writing – review and editing, formal analysis. **Ira Caspari-Gnann**: investigation, resources, writing – review and editing, project administration. **Nicole Graulich**: supervision, conceptualisation, methodology, writing – review and editing, project administration

Conflicts of interest

There are no conflicts to declare.

Appendix 1: exemplary arguments for the eight alternative reaction pathways

Fig. 6 shows exemplary arguments for the eight alternative reaction pathways of the diagnostic scaffold and adapted scaffolds. The arguments are exemplary and can be broadly extended. Moreover, arguments can be doubled because arguments for a plausible alternative reaction pathway are also arguments to justify why a certain alternative reaction pathway is implausible.

Appendix 2: an example of a science-related argument of the tasks on argumentation patterns

Fig. 7 shows an example of a science-related argument of the tasks on argumentation patterns. Students received tasks on argumentation patterns which varied in the number of arguments depending on their adapted scaffold. In the diagnostic scaffold, all groups received three arguments whereas, in the adapted scaffold, the ArgS and ArgConS groups received eight arguments and the ConS and ReaS groups received three arguments. All arguments were science-related. Students had to assign the argument components (claim, evidence, and

	Claim	Evidence	Reasoning	I don't know
Canaries serve as an "early warning system" in mining	●	○	○	○
because they can give miners an early indication of whether carbon monoxide is in the air	○	●	○	○
due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning.	○	○	●	○

Fig. 7 Example task on argumentation patterns. The correct answers are blue colour-coded.

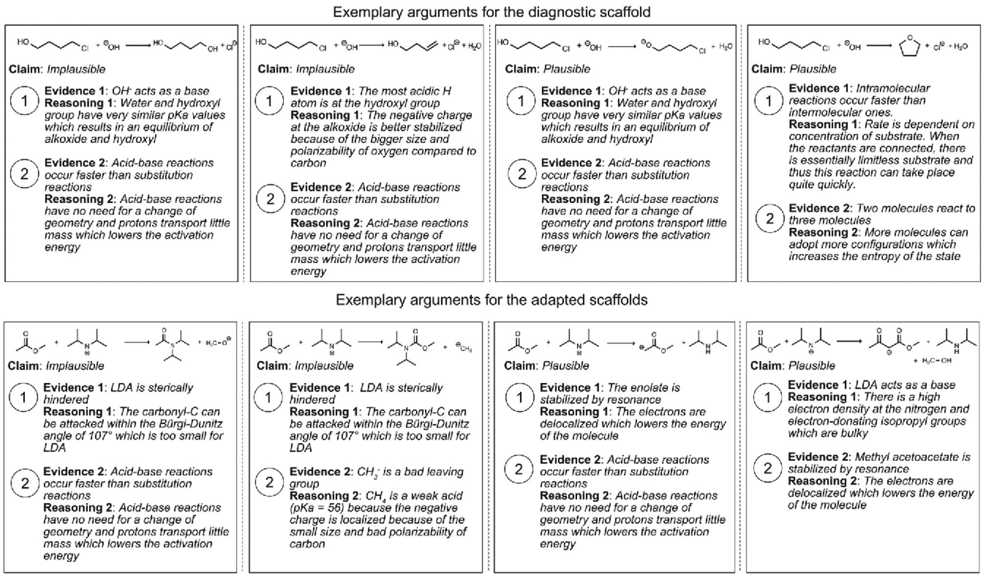


Fig. 6 Exemplary arguments for the four alternative reaction pathways in the diagnostic scaffold and the four alternative reaction pathways in the adapted scaffolds.

reasoning) to sentence components. They could always indicate that they do not know the answer.

Appendix 3: questions on chemical concepts with full-point student examples

Table 3 outlines the eight questions on chemical concepts students received in the diagnostic scaffold and adapted scaffolds. For each question, there is an example of a student's answer which was awarded full points in the scoring of concept knowledge. The student examples for the question on nucleophilicity and electrophilicity as well as on acidity and basicity are bipartite because these aspects were considered separately in the scoring system.

Appendix 4: task for building arguments for alternative reaction pathways

Fig. 8 illustrates the task of building arguments for alternative pathways for the formation of THF. Students judged the

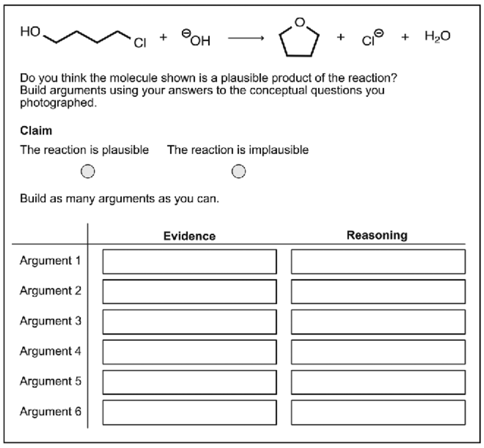


Fig. 8 An example of an alternative reaction pathway for the reaction of 4-chlorobutanol and hydroxide.

Table 3 Illustration of the questions on chemical concepts and student examples for each question, which received full points

Questions on chemical concepts	Student example
Decide whether steric aspects need to be considered in the reaction and explain why you think so.	*"I do not think that steric aspects need to be considered in the reaction because the reaction is taking place on a primary alkyl halide. A primary alkyl halide only has one other non-hydrogen substituent so it is relatively unhindered. Therefore, the –OH can attack the carbon without experiencing significant hindrance from other substituents which would be steric considerations."*
Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule *x* has the highest and molecule *y* the lowest pK_a)	*"I'd say the chlorobutanol is of a pK_a near 15 because I think that's the pK_a of water. The –OH itself probably has a pK_a of water too because the –OH is the conjugate base of water."*
Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.	*"The O of the OH group on the alkane acts as a nucleophile along with the O on the hydroxide ion because they have extra electron density."* *"I would expect the carbon bonded to the chlorine to be the most electrophilic site because chlorine is very electronegative and will pull electron density away from the carbon making it electrophilic. The carbon bonded to the hydroxyl group will be electrophilic for the same reason (oxygen is electronegative) but not as electronegative as the aforementioned carbon because oxygen is not as electronegative as chlorine."*
Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.	*"The 4-chloro-1-butanol reagent will act as a weak acid due to its mildly acidic hydroxyl group. This molecule will only lose its hydroxyl proton to moderately or strongly basic species that react to form a conjugate acid with a pK_a higher than that of 1-chloro-4-butanol (a higher pK_a corresponds to a less acidic, and thereby lower-energy, product)."* *"On the hydroxide, oxygen molecule is primary source of basicity due to negative charge on it."*
Determine whether you think there are any effects that stabilise your product compared to the reactants. If so, explain how the effect/s stabilise the product.	*"The product is stable because it is a five membered ring. This structure allows for optimal bond angles for sp^3 hybridization. The hydroxyl becomes protonated to form water, which is more stable since there are no formal charges in the molecule as oxygen forms two bonds."*
Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.	*"I do not think there are entropic effects that influence the reaction. There are two starting molecules and two products."*
Determine whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.	*"I think that the only electronic effect here is electronegativity and induction, as the Cl–C bond is polarized so that the carbon is a slightly positive center; there are no double bonds to induce resonance. The Cl^- is a better leaving group than OH partially because Cl is more electronegative than O, at least it is more electronegative towards other potential electrons than the effective electronegativity of an O bonded already to one H. The stronger inductive effects and electronegativity of Cl make it a better leaving group than the OH on the alcohol/chloride alkane."*
Decide whether the reaction is reactant- or product-favoured from an energetic perspective (enthalpy). Explain your thinking.	*"Product-favored. C–O bonds are stronger than C–Cl bonds due to the shorter length of C–O, so this substitution lowers the energy of the system."*

Table 4 Results of the Wilcoxon signed-rank test with Bonferroni-adjusted p-values for pre-post comparisons of the argumentation score and concept knowledge score. The correlation coefficient r was reported as effect size when p-values were significant. Significant results are highlighted in bold

Groups	M_{pre}	M_{post}	p	r
Argumentation score				
ArgS	10	14.5	**0.007**	0.68
ConS	13	14	0.823	
ArgConS	8.5	13.5	**<0.001**	0.74
ReaS	14	16	0.534	
Concept knowledge score				
ArgS	21	20	0.299	
ConS	16	19	**0.001**	0.66
ArgConS	13.5	18	**0.011**	0.57
ReaS	21	25	0.118	

plausibility of four alternative reaction pathways. Thereby, they created a claim and built evidence and reasoning statements to support and justify their decision. The text boxes were modified in size according to the number of words in the answers.

Appendix 5: summary of the results of the Wilcoxon signed-rank test

Table 4 shows the results of the Wilcoxon signed-rank test with Bonferroni-adjusted p-values for pre-post comparisons of the argumentation score and concept knowledge score.

Acknowledgements

This publication is part of the first author's doctoral (Dr rer. nat.) thesis at the Faculty of Biology and Chemistry, Justus-Liebig-University Giessen, Germany. We are thankful for all students willing to participate in the study and the professor for the opportunity to work with his students. Moreover, we thank all members of the Graulich group, Sascha Bernholt, Clay Bennett and especially David Kranz and Axel Langner for fruitful discussions and eye-opening moments. Leonie Sabine Lieber especially thanks the Verband der Chemischen Industrie (German Chemical Industry Association) for supporting her with the Kekulé Fellowship.

References

Azevedo R., Cromley J. G. and Seibert D., (2004), Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia?, *Contemp. Educ. Psychol.*, **29**(3), 344–370.

Azevedo R. and Gasevic D., (2019), Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges, *Comput. Hum. Behav.*, **96**, 207–210.

Belland B. R., (2017), *Instructional scaffolding in STEM education: Strategies and efficacy evidence*, Springer Nature.

Carle M. S. and Flynn A. B., (2020), Essential learning outcomes for delocalization (resonance) concepts: How are they taught, practiced, and assessed in organic chemistry?, *Chem. Educ. Res. Pract.*, **21**(2), 622–637.

Caspari I. and Graulich N., (2019), Scaffolding the structure of organic chemistry students' multivariate comparative mechanistic reasoning, *Int. J. Phys. Chem. Educ.*, **11**(2), 31–43.

Chen C. H., (2014), An adaptive scaffolding e-learning system for middle school students' physics learning, *Aust. J. Educ. Technol.*, **30**(3), 342–355.

Choi A., Hand B. and Greenbowe T., (2013), Students' Written Arguments in General Chemistry Laboratory Investigations, *Res. Sci. Educ.*, **43**(5), 1763–1783.

Cohen J., (1992), A power primer, *Psychol. Bull.*, **112**(1), 155–159.

Coppola B. P. and Pontrello J. K., (2014), Using errors to teach through a two-staged, structured review: Peer-reviewed quizzes and "What's wrong with me?", *J. Chem. Educ.*, **91**(12), 2148–2154.

Cruz-Ramirez de Arellano D. and Towns M. H., (2014), Students' understanding of alkyl halide reactions in undergraduate organic chemistry, *Chem. Educ. Res. Pract.*, **15**(4), 501–515.

Deng J. M. and Flynn A. B., (2021), Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items, *Chem. Educ. Res. Pract.*, **22**(3), 749–771.

Dood A. J., Dood J. C., de Arellano D. C. R., Fields K. B. and Raker J. R., (2020), Analyzing explanations of substitution reactions using lexical analysis and logistic regression techniques, *Chem. Educ. Res. Pract.*, **21**(1), 267–286.

Driver R., Asoko H., Leach J., Scott P. and Mortimer E., (1994), Constructing scientific knowledge in the classroom, *Educ. Res.*, **23**(7), 5–12.

Erduran S., (2007), *Argumentation in Science Education*, Springer, pp. 47–69.

Fan Y. C., Wang T. H. and Wang K. H., (2020), Studying the effectiveness of an online argumentation model for improving undergraduate students' argumentation ability, *J. Comput. Assist. Learn.*, **36**(4), 526–539.

Field A., Miles J. and Field Z., (2012), *Discovering statistics using R*, Sage publications.

Flynn A. B., (2021), *Problems and Problem Solving in Chemistry Education: Analysing Data, Looking for Patterns and Making Deductions*, pp. 145–165.

Graulich N. and Caspari I., (2021), Designing a scaffold for mechanistic reasoning in organic chemistry, *Chem. Teach. Int.*, **3**(1), 19–30.

Hogan K. and Maglienti M., (2001), Comparing the epistemological underpinnings of students' and scientists' reasoning about conclusions, *J. Res. Sci. Teach.*, **38**(6), 663–687.

Kang H., Thompson J. and Windschitl M., (2014), Creating Opportunities for Students to Show What They Know: The Role of Scaffolding in Assessment Tasks, *Sci. Educ.*, **98**(4), 674–704.

Lajoie S. P., (2005), Extending the scaffolding metaphor, *Instr. Sci.*, **33**(5–6), 541–557.

Lazarou D. and Erduran S., (2020), "Evaluate What I Was Taught, Not What You Expected Me to Know": Evaluating Students' Arguments Based on Science Teachers' Adaptations to Toulmin's Argument Pattern, *J. Sci. Teach. Educ.*, **32**(3), 306–324.

Lee H. S., Gweon G. H., Lord T., Paessel N., Pallant A. and Pryputniewicz S., (2021), Machine Learning-Enabled

Automated Feedback: Supporting Students' Revision of Scientific Arguments Based on Data Drawn from Simulation, *J. Sci. Educ. Technol.*, **30**(2), 168–192.

Lieber L. and Graulich N., (2020), Thinking in Alternatives-A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry, *J. Chem. Educ.*, **97**(10), 3731–3738.

Lieber L. and Graulich N., (2022), Investigating Students' Argumentation when Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry, *Chem. Educ. Res. Pract.*, **23**, 38–54.

Luo X. L., Wei B., Shi M. and Xiao X., (2020), Exploring the impact of the reasoning flow scaffold (RFS) on students' scientific argumentation: based on the structure of observed learning outcomes (SOLO) taxonomy, *Chem. Educ. Res. Pract.*, **21**(4), 1083–1094.

McGinn M. K. and Roth W.-M., (1999), Preparing students for competent scientific practice: Implications of recent research in science and technology studies, *Educ. Res.*, **28**(3), 14–24.

McNeill K. L. and Krajcik J., (2012), *Book study facilitator's guide: Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*, New York: Pearson Allyn & Bacon.

McNeill K. L., Lizotte D. J., Krajcik J. and Marx R. W., (2006), Supporting students' construction of scientific explanations by fading scaffolds in instructional materials, *J. Learn. Sci.*, **15**(2), 153–191.

Milkman K. L., Chugh D. and Bazerman M. H., (2009), How can decision making be improved?, *Perspect. Psychol. Sci.*, **4**(4), 379–383.

Moon A., Stanford C., Cole R. and Towns M., (2016), The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry, *Chem. Educ. Res. Pract.*, **17**(2), 353–364.

Noroozi O., Kirschner P. A., Biemans H. J. A. and Mulder M., (2017), Promoting Argumentation Competence: Extending from First- to Second-Order Scaffolding Through Adaptive Fading, *Educ. Psychol.*, **30**(1), 153–176.

Osborne J. F. and Patterson A., (2011), Scientific argument and explanation: A necessary distinction?, *Sci. Educ.*, **95**(4), 627–638.

Pabuccu A. and Erduran S., (2017), Beyond rote learning in organic chemistry: the infusion and impact of argumentation in tertiary education, *Int. J. Sci. Educ.*, **39**(9), 1154–1172.

Pea R. D., (2004), The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity, *J. Learn. Sci.*, **13**(3), 423–451.

Petritis S. J., Kelley C. and Talanquer V., (2021), Exploring the impact of the framing of a laboratory experiment on the nature of student argumentation, *Chem. Educ. Res. Pract.*, **22**(1), 105–121.

Petritis S. J., Kelley C. and Talanquer V., (2022), Analysis of factors that affect the nature and quality of student laboratory argumentation, *Chem. Educ. Res. Pract.*, **23**, 257–274.

Popova M. and Bretz S. L., (2018), "It's Only the Major Product That We Care About in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams, *J. Chem. Educ.*, **95**(7), 1086–1093.

Rosenthal R., (1991), *Meta-analytic procedures for social research*, Newbury Park, CA: Sage.

Sadler T. D., (2004), Informal reasoning regarding socioscientific issues: A critical review of research, *J. Res. Sci. Teach.*, **41**(5), 513–536.

Sandoval W. A. and Millwood K. A., (2005), The quality of students' use of evidence in written scientific explanations, *Cogn. Instr.*, **23**(1), 23–55.

Shute V. J. and Zapata-Rivera D., (2008), *Handbook of research on educational communications and technology*, New York, NY: Taylor and Francis, pp. 277–294.

Smith M. K., Wood W. B., Adams W. K., Wieman C., Knight J. K., Guild N. and Su T. T., (2009), Why peer discussion improves student performance on in-class concept questions, *Science*, **323**(5910), 122–124.

Songer N. B. and Gotwals A. W., (2012), Guiding explanation construction by children at the entry points of learning progressions, *J. Res. Sci. Teach.*, **49**(2), 141–165.

Stowe R. L., Scharlott L. J., Ralph V. R., Becker N. M. and Cooper M. M., (2021), You Are What You Assess: The Case for Emphasizing Chemistry on Chemistry Assessments, *J. Chem. Educ.*, **98**(8), 2490–2495.

Toulmin S. E., (2003), *The Uses of Argument*, Cambridge University Press.

Van de Pol J., Volman M. and Beishuizen J., (2010), Scaffolding in teacher–student interaction: A decade of research, *Educ. Psychol.*, **22**(3), 271–296.

Van Eemeren F. H., Garssen B., Krabbe E. C. W., Snoeck Henkemans A. F., Verheij B. and Wagemans J. H. M., (2014), *Handbook of Argumentation Theory*, Dordrecht: Springer, pp. 203–256.

Vygotsky L. S., (1980), *Mind in society: The development of higher psychological processes*, Harvard University Press.

Walker J. P., Van Duzor A. G. and Lower M. A., (2019), Facilitating Argumentation in the Laboratory: The Challenges of Claim Change and Justification by Theory, *J. Chem. Educ.*, **96**(3), 435–444.

Walqui A., (2006), Scaffolding instruction for English language learners: A conceptual framework, *Int. J. Biling. Educ. Biling.*, **9**(2), 159–180.

Watts F. M., Zaimi I., Kranz D., Graulich N. and Shultz G. V., (2021), Investigating students' reasoning over time for case comparisons of acyl transfer reaction mechanisms, *Chem. Educ. Res. Pract.*, **22**(2), 364–381.

Wecker C. and Fischer F., (2014), Where is the evidence? A meta-analysis on the role of argumentation for the acquisition of domain-specific knowledge in computer-supported collaborative learning, *Comput. Educ.*, **75**, 218–228.

Wood D., Bruner J. S. and Ross G., (1976), The role of tutoring in problem solving, *J. Child Psychol. Psychiatry*, **17**(2), 89–100.

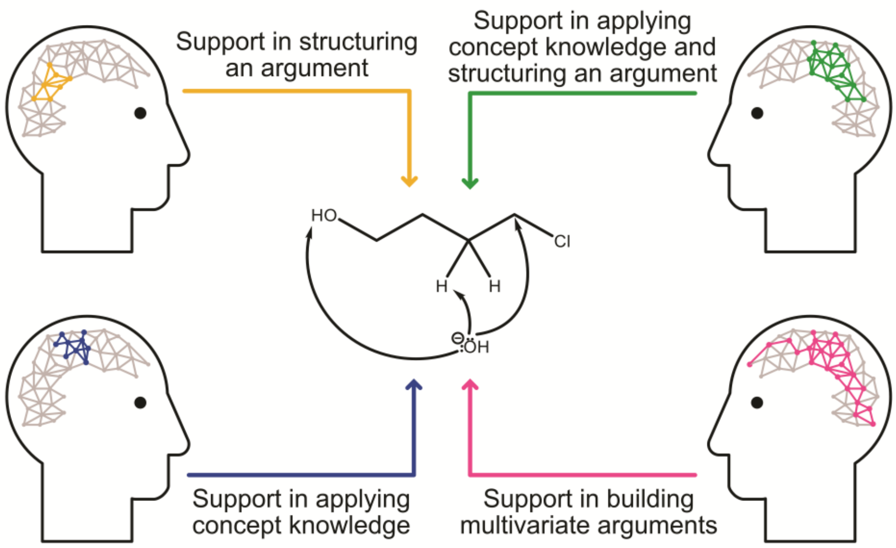
Yik B. J., Dood A. J., de Arellano D. C. R., Fields K. B. and Raker J. R., (2021), Development of a machine learning-based tool to evaluate correct Lewis acid-base model use in written responses to open-ended formative assessment items, *Chem. Educ. Res. Pract.*, **22**(4), 866–885.

Yuriev E., Naidu S., Schembri L. S. and Short J. L., (2017), Scaffolding the development of problem-solving skills in chemistry: guiding novice students out of dead ends and false starts, *Chem. Educ. Res. Pract.*, **18**(3), 486–504.

Zhou Z. H., (2016), Learnware: on the future of machine learning, *Front. Comput. Sci.*, **10**(4), 589–590.

5 Students' Individual Needs Matter – A Training to Adaptively Address Students' Argumentation Skills in Organic Chemistry

In Chapter 5, the original paper is represented. It is reproduced from Lieber, Ibraj, Caspari-Gnann and Graulich (2022) with permission of the American Chemical Society.



Lieber, L., Ibraj, K., Caspari-Gnann, I. & Graulich, N. (2022). Students' individual needs matter – a training to adaptively address students' argumentation skills in organic chemistry, *Journal of Chemical Education, 99*(7), 2754-2761.

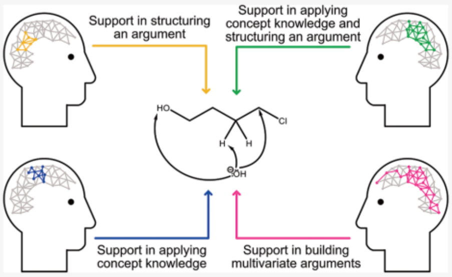
JOURNAL OF CHEMICAL EDUCATION

pubs.acs.org/jchemeduc Activity

Students' Individual Needs Matter: A Training to Adaptively Address Students' Argumentation Skills in Organic Chemistry

Leonie Sabine Lieber, Krenare Ibraj, Ira Caspari-Gnann, and Nicole Graulich*

Cite This: *J. Chem. Educ.* 2022, 99, 2754−2761 Read Online



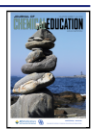
ABSTRACT: In chemistry, building arguments and applying concept knowledge is closely linked to evaluating claims, supporting claims with evidence, and justifying the linkage of evidence to claim with reasoning. However, previous studies revealed that, when building arguments, students experience challenges either in differentiating between argument components, in applying concept knowledge, or in building multivariate arguments. Scaffolding may remediate these challenges by supporting students as they focus on the expected argument structure and/or activate the required concept knowledge. As students enter the classroom with different prior knowledge, supports need to be adapted to students' needs. Thus, we designed a two-part argumentation training. The first part is a diagnostic training, in which students receive training for building arguments while their performance is analyzed. The second part consists of four trainings, adapted to the area in which each student experienced the greatest challenges, e.g., (1) in differentiating between argument components, (2) in applying concept knowledge, (3) in both areas, or (4) in building multivariate arguments. The tasks in the trainings center on building arguments on alternative reaction pathways in organic chemistry and combine a multitude of chemical concepts, such as nucleophilicity, basicity, enthalpy, or entropy. There were 64 students enrolled in an Organic Chemistry II course who participated in the training. Evaluation of the two-part training revealed (a) the effectiveness of the training and (b) how students evaluated the training themselves.

KEYWORDS: *Upper-Division Undergraduate, Organic Chemistry, Problem Solving/Decision Making, Internet/Web-Based Learning, Mechanisms of Reactions*

■ INTRODUCTION

Building arguments is one of the most important skills a chemist needs to gain expertise in since it is necessary to critically engage with hypotheses and concepts, or to propose novel solutions.[1−4] Therefore, argumentation occurs whenever decisions and judgements are made,[5] which is why argumentation is inextricably linked to science and has been a central component of science education research for decades.[6−8] However, by highlighting the importance of argumentation, students may pose the question regarding what scientific argumentation is about.[9] In the process of building scientific arguments, concept knowledge is linked and applied to evaluate claims and weigh evidence.[10] Argumentation and concept knowledge are closely linked since students can develop a deeper conceptual understanding when building arguments, but they also need concept knowledge to be able to build arguments.[8,11,12] By applying concept knowledge when building arguments, students are often required to build multivariate arguments which means that they coordinate several chemical concepts to back up their arguments.[13,14] Thus, it is equally important to support students in learning scientific concepts and scientific argumentation.[15] However, students may experience challenges in the appropriate use of argument components (i.e., claim, evidence, and reasoning)[16−20] and in building chemically sound and multivariate arguments.[13,20−25] Several instructional approaches have already been designed in chemistry education to support students in building arguments and to help researchers to better understand how students built arguments in different contexts and activate their concept knowledge.[26−28] For example, a recently designed instructional approach by Petritis et al. consists of a claim−evidence−rationale scaffold for argumentation to characterize students' written arguments, which students constructed by evaluating results from an organic chemistry laboratory.[29] By characterizing arguments and attributing their findings to different laboratory factors, the authors could show that the integration of chemical concepts

Received: March 10, 2022
Revised: June 2, 2022
Published: June 24, 2022



https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

in students' arguments varied depending on the experiment, and that scaffolding was beneficial for certain students but that students need further support.[29] This example, as well as other studies,[22,30,31] indicates that argumentation needs to be adaptively scaffolded, tailored to students' individual needs.[32,33] Accordingly, this need may also occur when judging alternative reaction pathways.

Our previous studies revealed that students are typically not used to building arguments and that students experience challenges familiar in the literature, such as supporting claims with evidence and reasoning or appropriately applying chemical concepts in building arguments.[22,34] In order to provide students with an adapted argumentation training in organic chemistry, that acknowledges students' individual challenges, a two-part argumentation training was designed. In a first step, the individual challenges students experience when building arguments on alternative reaction pathways in organic chemistry were diagnosed, and in a second step, their challenges are addressed with four adapted trainings. The goal of the activity presented here is to outline this training and to demonstrate learning gains before and after the adapted trainings (i.e., students' performance and their ability to revise their initial claims), and students' evaluation of the training.

■ TWO-PART ARGUMENTATION TRAINING

To help students overcome their struggles in building arguments, a two-part training was designed in an online learning environment using Qualtrics. On the basis of previous studies,[22,34] it became clear that when building arguments, students experience challenges either (1) in using argument components (i.e., claim, evidence, and reasoning), (2) in using concept knowledge, (3) in both areas, or (4) in building multivariate arguments. Therefore, in the diagnostic training, not only are students supported in building arguments and using concept knowledge, but also students' answers are scored, based on a manual scoring system (see Figure 1 and Supporting Information) to reveal in which of the addressed areas students are struggling. The scoring system is in two-parts and scores concept knowledge and argumentation separately. Moreover, the scoring system encompasses definitions for the argument components, i.e., evidence and reasoning, and information regarding the answers to the questions on chemical concepts (e.g., "Determine at which positions you think the molecules react as a nucleophile", corresponding information for the reaction of the diagnostic scaffold "OH-group, OH^-, Cl"). Table 1 illustrates one example for each of the two parts of the scoring system. The scoring was conducted by the authors, in which they discussed each student answer and each built argument with the criteria of the scoring system. On the basis of their scoring results, students are then given one of four adapted trainings which support students in the area in which they experience the greatest difficulties.

Figure 1. Via the QR Code or the DOI 10.17605/OSF.IO/4ZPN9, the trainings and other materials are available online. The materials include file formats such as QSF, XML, DOCX, and Unicode Text in English and German.

Table 1. Exemplary Excerpts of the Scoring System for Concept Knowledge and Argumentation

Concept Knowledge Scoring		
Nucleophilicity	Empty field/"I do not know"	0
	Name the wrong nucleophilic centers	0
	Name one nucleophilic center correctly	1
	Name at least two nucleophilic centers correctly	2
Argumentation Scoring		
Evidence	Incorrectly formed pieces of evidence prevail (<50% correct)	0
	Correctly formed pieces of evidence prevail (≥50% correct)	1
	All pieces of evidence are formed correctly	2

The design of the trainings was primarily theory-driven. The claim−evidence−reasoning model was chosen as the basis of the argumentation training because it is applicable to our students' needs as they did not have experience on building arguments, and the claim−evidence−reasoning model is well-suited for starting argumentation in class.[35,36] Moreover, we chose to use four instructional approaches suggested by Kang et al. to design the training.[37] *Instructional modeling*, the first instructional approach, gives examples to students to emphasize what is expected from them. Thus, students are given examples in the trainings of how to use argument components in the science context.[37] By implementing *contextualization*, ideas are made accessible for students, which was applied in the trainings by structuring the argument in partial steps and by illustrating the linkage between argument components and concept knowledge.[38,39] This is supported by *bridging* because this instructional approach focuses on the connection of prior knowledge and new contexts.[40] Therefore, students are encouraged in the trainings to activate their prior knowledge by answering questions on multiple chemical concepts,[41] because it has been shown not only that students tend to integrate a single concept into the decision-making process instead of multiple factors[25,42,43] but also that students think more deeply about a reaction by including multiple reaction centers.[44] The last instructional approach is *developing metacognition*, which was applied to foster students' reflection process by assessing the difficulty, student confidence, and an evaluation of the tasks in the trainings.[37] The tasks on alternative reactions pathways used in the trainings are already published by Lieber and Graulich[34] and intend to cause a cognitive conflict in students' thinking[45,46] and thus more analytical thinking, making it more likely for students to integrate chemical concepts into their decision-making process.[47] The trainings and other Supporting Information can be accessed via the QR code in Figure 1 and the Supporting Information.

https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

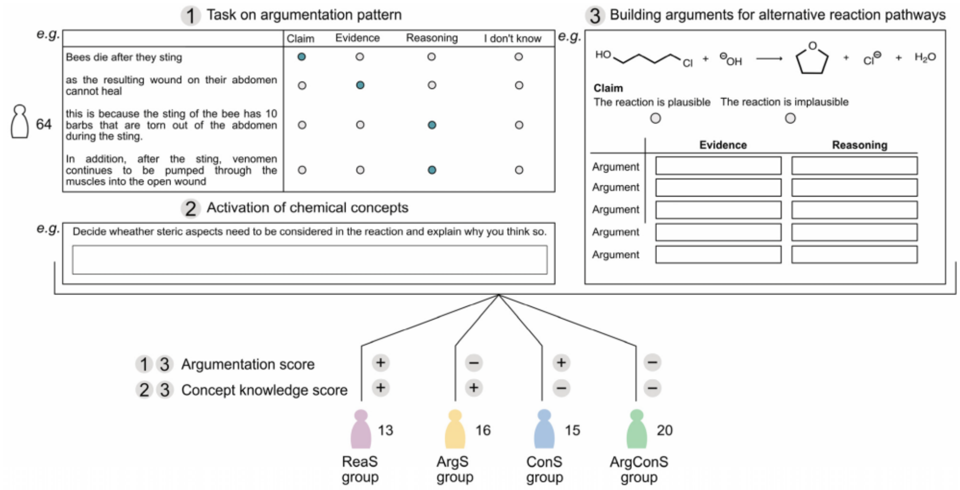


Figure 2. Exemplary illustration of diagnostic training tasks that were incorporated into the scoring and the grouping of students into the adapted training groups based on a high (+) or low (−) argumentation and concept knowledge score. Students of the ReaS group received support in building multivariate arguments in the adapted training, students of the ArgS group in building arguments, students of the ConS group in applying concept knowledge, and students of the ArgConS group in building arguments and applying concept knowledge.

Diagnostic Training

All students work on the same diagnostic training. In a first step, the students have to predict the product of the reaction of 4-chlorobutanol and hydroxide. In a second step, students first receive general explanations of the structure of an argument, followed by three exercises on building science-related arguments (see Figure 2, panel 1). The examples are chosen to emphasize the incompleteness of arguments, which only consist of evidence but not reasoning statements, but which represent typical student responses. Conversely, complete examples are shown in which a piece of evidence is justified with several reasoning statements. In a next step, students are prompted to think about relevant concept knowledge (see Figure 2, panel 2) that applies to the task at hand. There are 10 questions that address different chemical concepts and influential factors (e.g., nucleophilicity, entropy, or steric aspects) using the example of the reaction of 4-chlorobutanol and hydroxide to activate students' prior knowledge. In the next step, students consecutively judge the plausibility of four alternative reaction pathways for the reaction of 4-chlorobutanol and hydroxide (see Figure 2, panel 3). Here, the students first select a claim as to whether the reaction product is plausible or implausible. Students are then asked to build as many evidence and reasoning statements as they can to support their claim. By building arguments about the alternative reaction pathways, students experience that it is not always just a matter of predicting a main product, but that alternative products or byproducts must be considered.[48] Next, students can indicate whether they will keep their product, predicted at the beginning, or whether they will choose one or more of the alternative reaction products. Finally, students evaluate the training and their own performance.

After the diagnostic training, students' answers were scored, and an argumentation score (maximum 25 points) and a concept knowledge score (maximum 29 points) were determined for each student. From the maximum score, 65% was set as the threshold for assigning students in the adapted training groups, as this corresponds to the passing D grade at many US universities. The scoring system is provided via the QR code in Figure 1 and in the Supporting Information.

Adapted Trainings

On the basis of students' scoring results in the diagnostic training, they are assigned to one of four adapted trainings (see Figure 2). The structure of the adapted trainings is comparable to the structure of the diagnostic training. However, the adapted trainings use the reaction of methyl acetate and diisopropylamide. This reaction was chosen because both reactions can be discussed under the lens of the competition of nucleophilicity and basicity.

The four adapted training groups differ in the type of additional support students receive and are based on the area in which they experienced the greatest challenges with structuring and building arguments (ArgS), applying concept knowledge (ConS), building arguments and applying concept knowledge (ArgConS), or building multivariate arguments (ReaS).

Students in the ArgS group experience challenges in building arguments with argument components (i.e., claim, evidence, and reasoning). Students are assigned to this group when their argumentation score is below the threshold of 16 points and the concept knowledge score is above the threshold of 18 points. For fostering their argumentation skills, students receive eight science-related argumentation patterns in the adapted training. These include both complete and incomplete examples, as well as arguments that are justified with multiple reasoning statements to make the expectations of an appropriate argument transparent.[37,40] Additionally, students are given definitions for claim, evidence, and reasoning when building arguments on their own.

https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

Journal of Chemical Education pubs.acs.org/jchemeduc Activity

The second group is the ConS group. Students who score more than 16 points in the argumentation score and less than 18 points in the concept knowledge score are assigned to this group, which provides them with support in applying concept knowledge. Students are first asked to activate their prior knowledge by answering 10 questions on chemical concepts that apply to the task at hand. While building autonomous arguments, students are given the expected answers to these questions, but they still have to interpret them by themselves. This step aims at supporting students in noticing the necessary features.[49]

Students in the ArgConS group are supported in both building arguments and using concept knowledge. Students in the ArgConS group have less than 16 points in the argumentation score and less than 18 points in the concept knowledge score. For this reason, students receive a combination of the trainings of the ArgS group and the ConS group.

The fourth group is the ReaS group. Students in this group exceed the threshold of 16 points (argumentation score) and 18 points (concept knowledge score). Those students already achieve good results in the diagnostic training; in the adapted training, they are challenged to further enhance their argumentation skills. Therefore, at the beginning of the ReaS training, students are given three multivariate argumentation patterns. While building autonomous arguments, students are encouraged to build multivariate arguments, i.e., build three reasoning statements per piece of evidence.

Implementation

The implementation of the two-part argumentation training took place in an Organic Chemistry II course at a university in the northeastern United States in April and May 2021 using the application Qualtrics. Students had 48 h to complete each training, whereas the completion time was between 30 and 90 min. It was ensured with the teacher of the course that all topics of the trainings were covered and discussed previously in the course. There were 64 students who participated in the study for which they received extra credit points (1% of their total course grade). The demographic data are shown in Table 2. The diagnostic training took place in the 10th week, and the adapted trainings took place in the 13th week of the semester. The first and second authors scored students' answers within 1 week each for the diagnostic training and the adapted trainings, respectively. Students' answers were sometimes short and not very precise. Therefore, students' answers were taken verbatim to ensure that the manual scoring was consistent and objective.

Table 2. Demographic Data of the 64 Students Who Participated in the Two-Part Argumentation Training

Two-Part Argumentation Training					
Gender					
Female		Male		Nonbinary	
34		29		1	
Race					
White	Asian	Black	More than others	Latino/a/x	Others
36	20	3	3	1	1
Student Major					
Biochemistry	Chemistry	Biology	Others		
31	14	8	11		

All students created a user code as a pseudonym, which does not reflect their race, gender, or other identities.

■ STUDENTS' LEARNING GAINS

After the diagnostic training and after the adapted trainings, students' answers were scored each time to ascertain an argumentation score and a concept knowledge score (see the QR Code in Figure 1 and the Supporting Information for the scoring system). To determine an improvement in students' performance after the two trainings, Wilcoxon signed-rank tests were performed. The correlation coefficient r was reported as effect size, which was defined as the following: $0.10 \leq r \leq 0.30$ as small effect, $0.30 \leq r \leq 0.50$ as medium, and $r \geq 0.50$ as large.[50] The analysis revealed that the students improved their performance in building arguments significantly with a medium effect ($M_{DS} = 11.5$; $M_{AS} = 14$; $V = 1271$; $p = < 0.001$; $r = 0.48$). Moreover, students improved their use of concept knowledge significantly with a medium effect ($M_{DS} = 17$; $M_{AS} = 19$; $V = 1303, 5$; $p = 0.002$; $r = 0.48$) (see Figure 3). A detailed description of each adapted training groups' performance improvement can be found in a more detailed publication by Lieber et al.[51] The results indicate that the adapted trainings helped students to improve the building of arguments and supporting arguments with concept knowledge. However, more research needs to be done to ascertain if students would retain their argumentative skills after receiving this training.

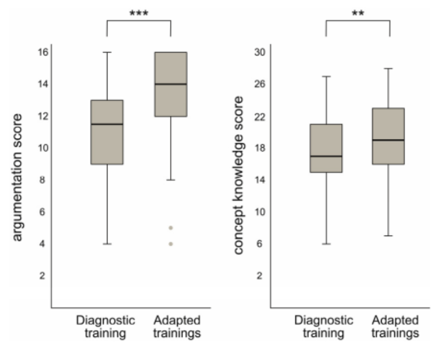


Figure 3. Argumentation score and concept knowledge score for all students after the diagnostic training and the adapted trainings. Horizontal stripes in the box plots indicate median-values. Significance levels of the comparisons are indicated (NS. $p > 0.05$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

A second step to investigate students' learning gains is the analysis of students revising their initial claims after building arguments with the use of concept knowledge. At the beginning of each training (i.e., diagnostic and adapted training), the students were asked to make a claim about the product formed during the reaction given. After building arguments, they had the chance to revise this initial claim and to potentially correct it. Figure 4 summarizes students' "most plausible" claims separated for each adapted training group and with regard to the diagnostic and adapted training, respectively. Thus, percentages are given in ranges because the percentages differ within the four groups. All reaction products that were claimed as "most plausible" were aggregated and represent a total of 100%. In the beginning of the diagnostic training, at

https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

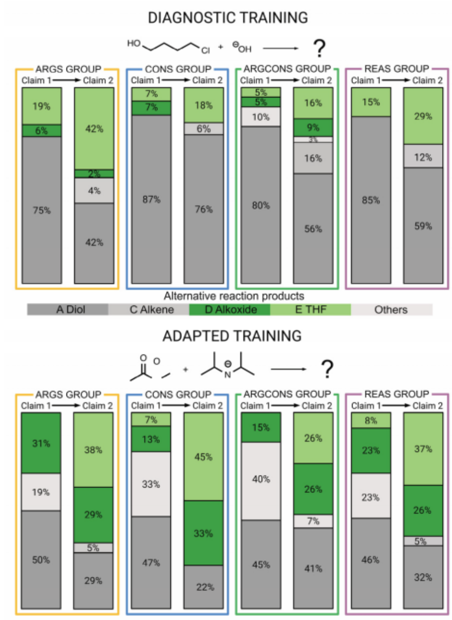


Figure 4. Total of students' "most plausible" claims made in the diagnostic training and the adapted trainings separated for each of the four training groups. The correct products are illustrated in green. "Others" refers to incorrect product such as aldehydes or intermediates. Students were able to claim more than one product as "most plausible".

least three-quarters of the students claimed the diol as the reaction product (75%−87%), whereas only 5%−19% of the products were claimed as the correct product THF in claim 1. However, after all students received support for building arguments and applied their concept knowledge, many students revised their initial claims at the end of the diagnostic training (claim 2). Thereby, the percentage of the correct product and its precursor (THF and alkoxide) increased and resulted in a total of 18%−44% (see Figure 4, upper half).

In the adapted trainings, all groups received the same organic chemistry reaction of methyl acetate and diisopropylamide. No students of the ArgS group and ArgConS group have claimed the product methyl acetoacetate to be plausible at the beginning of the adapted training, but several students of all groups have already formed the correct precursor (13%−31%) (see Figure 4, lower half). After building arguments for alternative reaction pathways, students of all groups revised their initial claims and, thus, increased the claims of the correct product and its precursor (methyl acetoacetate and enolate) greatly, which resulted in a total percentage of 52%−78%. These findings reveal that through the training students are engaged in successfully revising their initial claim. However, compared to a previously published interview study,[22] it became apparent that students are less likely to revise their claims in written arguments compared to building arguments verbally, as already documented in the chemistry education literature.[3,18,52,53] Reasons for sticking with an erroneous claim may be that it is difficult to change a person's point of view in such a short period of time,[3,53] that students perceive a claim change as a sign of failure of their previous work,[18] or that students experience challenges in applying their concept knowledge.[3] Especially in the diagnostic training, many students maintained their erroneous claims as students did not realize that hydroxide acts as a base and not as a nucleophile in the reaction of 4-chlorobutanol and hydroxide. Nevertheless, the adapted training has an impact on students to propose the correct reaction products.

■ STUDENTS' FEEDBACK

In addition to the analysis of students' learning gains, we were also interested in how the students evaluated the trainings. After arguing for or against each alternative reaction pathway, students rated the difficulty of the task and their confidence to

	Difficulty								Confidence							
	ArgS group		ConS group		ArgConS group		ReaS group		ArgS group		ConS group		ArgConS group		ReaS group	
Sub-task	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD	$\bar{x}$	SD
1.1	2.44	0.79	2.33	0.79	2.50	0.50	2.00	0.68	2.50	0.61	2.20	0.91	2.40	0.66	2.08	0.73
1.2	2.31	0.85	2.53	0.62	2.45	0.67	2.31	0.82	2.31	0.77	2.47	0.72	2.30	0.64	2.23	0.58
1.3	2.25	0.66	2.53	0.72	2.55	0.59	2.54	0.84	2.25	0.66	2.53	0.72	2.25	0.77	2.46	0.75
1.4	2.31	0.66	2.73	0.57	2.85	0.85	2.54	0.75	2.56	0.58	2.53	0.72	2.90	0.77	2.54	0.75
Total	2.31	0.75	2.53	0.69	2.59	0.68	2.35	0.81	2.34	0.67	2.43	0.78	2.46	0.76	2.33	0.73
2.1	2.46	0.70	2.67	0.47	2.90	0.44	2.38	0.62	2.50	0.61	2.47	0.50	2.70	0.64	2.23	0.42
2.2	2.13	0.70	2.53	0.72	2.65	0.73	2.00	0.55	2.00	0.79	2.20	0.75	2.45	0.74	1.77	0.58
2.3	2.19	0.63	2.45	0.65	2.45	0.74	2.15	0.77	2.31	0.58	2.53	0.62	2.40	0.80	2.23	0.80
2.4	2.69	0.77	3.07	0.68	2.95	0.74	2.62	0.74	2.31	0.68	2.73	0.85	3.05	0.80	2.54	0.84
Total	2.39	0.74	2.77	0.67	2.74	0.70	2.29	0.72	2.28	0.70	2.48	0.72	2.65	0.79	2.19	0.73

Figure 5. Results of students' rating of the tasks' difficulty and their confidence to build arguments after the diagnostic training (1.1−1.4) and after the adapted trainings (2.1−2.4) using a Likert scale. The scale for response ranged from 1 to 4: difficulty, 1 = easy, 2 = rather easy, 3 = rather difficult, 4 = difficult; confidence, 1 = very confident, 2 = rather confident, 3 = rather unconfident, 4 = unconfident; $\bar{x}$ = means, SD = standard deviation.

https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

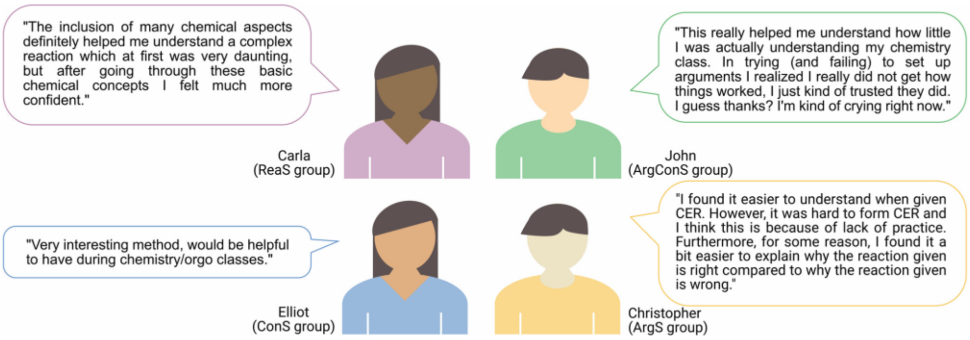


Figure 6. Examples of students' written feedback given in the trainings.

build arguments, each of which consisted of a Likert scale from 1 to 4 (see Figure 5). The results indicate that students rated the difficulty and confidence to build arguments comparably, so that each alternative reaction pathway was rated as rather easy and their confidence as rather confident in both the diagnostic training and the adapted training. This illustrates that the tasks were not too easy or too difficult so that students might feel overwhelmed. Students in all four groups were thus within the desired range in which they still felt a challenge despite support. Additionally, students in the ArgS group and ReaS group rated the tasks easier on average (2.31 and 2.35) than the students in the ConS group and ArgConS group (2.53 and 2.59). Since the last two groups experienced challenges in the diagnostic training in terms of concept knowledge, it can be assumed that the use of concept knowledge has a greater influence on the subjective assessment of difficulty than building arguments.

Students also had the opportunity to express individual feedback. Carla is a student from the ReaS group (see Figure 6). She expressed that the reactions were complex for her and that she was intimidated. However, by prompting her to apply her concept knowledge when building arguments, she not only felt more encouraged but also was strengthened in her understanding. Carla's feedback summarizes what the argumentation training intended: supporting students to build chemically sound arguments without telling them the solution process. John's quote, as a student from the ArgConS group, illustrates that the training can lead to students being pushed to their limits. By working with the training, John realized that while he thought he understood organic chemistry, when building arguments, he became aware that he struggled with concept knowledge. The adapted training was designed to encourage students like John to apply concept knowledge when building arguments. Thereby, it is possible for students to become explicitly aware that solving organic chemistry reactions consists of applying chemical concepts, which may cause students to feel unsettled. Elliot's feedback, a student from the ConS group, for instance, calls for more opportunity of this type of training in organic chemistry. This seems to be surprising because the literature revealed that students often resist when being introduced to a new learning environment and demand a traditional learning setting,[54,55] which may be, besides other aspects, a reason why traditional learning settings are omnipresent. However, it is encouraging that students ask for the use of the training in traditional learning settings to support them in applying concept knowledge and building arguments. Christopher, a student of the ArgS group, expressed that he found it difficult to build claim, evidence, and reasoning and that is was easier to build arguments for plausible reaction products than for implausible ones. This comment is not surprising as students typically learn the formation of main products but are often unaware of possible byproducts.[22,48] Regular use of the training can not only familiarize students with the fact that chemical reactions form byproducts but also give them practice in building and differentiating between argument components.

CONCLUSION

This activity described herein illustrates a two-part argumentation training that is able to support students in building arguments and using concept knowledge to improve the quality of their arguments. The training diagnoses in a first step which challenges students' experience in building arguments and supports them in a second step in the respective areas (i.e., argument components and concept knowledge) where they experienced the greatest challenges. Looking at students' learning gains, it is evident that students improved significantly in both the use of argument components (i.e., claim, evidence, and reasoning) and in the use of concept knowledge. The linkage between argumentation and concept knowledge is mentioned in the literature[56] and is considered a key aspect regarding the quality of an argument.[38,39] Adapting the training in the second part to students' needs seems to positively influence students' decision-making process. Additionally, after building arguments in both trainings, students were often able to revise their initial claim and form the correct reaction product. Students also indicated that they appreciated the two-part argumentation training, used more chemical concepts, and would like to see it used in class.

Using the argumentation training allows teachers to provide differentiated support not only for students who encounter difficulties, but also for already high-achieving students to further deepen their ability to build arguments. Thereby, the trainings can be applied in class in multiple ways. First, the training can be fully used, as illustrated herein with the inclusion of the scoring, but it is also possible to use or discuss only individual aspects in class, e.g., the appropriate use of argument components by using the exercises on argumentation patterns. Moreover, students could choose by themselves in which area they want to receive additional training or the teacher decides that specific aspects of the trainings can be part of, for example, classroom discussions. Thereby, it is not mandatory to score students' answers to implement aspects of the training which makes it suitable on a larger scale in the classroom. Moreover, the argumentation on alternative reaction pathways can be adapted by using different reactions since the chemical concepts can be applied to any reaction in organic chemistry. These options make the argumentation training suitable for both diagnosing and fostering students in building arguments.

However, scoring students' arguments is time-consuming, especially in classrooms with a great number of students. To face this challenge, future research will be focused on an automatic scoring system to support teachers in applying the argumentation training.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.2c00213.

Two-part argumentation training and scoring systems of the diagnostic training and the adapted training (PDF, DOCX)

■ AUTHOR INFORMATION

Corresponding Author

Nicole Graulich − *Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen 35392, Germany*; orcid.org/0000-0002-0444-8609; Email: Nicole.Graulich@dc.jlug.de

Authors

Leonie Sabine Lieber − *Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen 35392, Germany*; orcid.org/0000-0001-9955-6655

Krenare Ibraj − *Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen 35392, Germany*; orcid.org/0000-0002-3904-011X

Ira Caspari-Gnann − *Department of Chemistry, Tufts University, Medford, Massachusetts 02155, United States*; orcid.org/0000-0003-1728-8656

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.2c00213

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This publication is part of the first author's doctoral thesis (Dr. rer. nat.) at the Faculty of Biology and Chemistry, Justus-Liebig-University Giessen, Germany. We thank all students for their participation and the professor for working with his students. L.S.L. thanks especially the Verband der Chemischen Industrie (German Chemical Industry Association) for supporting her with the Kekulé Fellowship.

■ REFERENCES

(1) Council, N. R. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; The National Academies Press, 2012. DOI: 10.17226/13165.

(2) Longino, H. E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*; Princeton University Press, 1990. DOI: 10.1515/9780691209753.

(3) Novak, A. M.; Treagust, D. F. Adjusting claims as new evidence emerges: Do students incorporate new evidence into their scientific explanations? *J. Res. Sci. Teach.* **2018**, *55* (4), 526−549.

(4) Osborne, J. F.; Patterson, A. Scientific argument and explanation: A necessary distinction? *Sci. Educ.* **2011**, *95* (4), 627−638.

(5) Kuhn, D. Thinking as argument. *Harv. Educ. Rev.* **1992**, *62* (2), 155−179.

(6) Erduran, S.; Ardac, D.; Yakmaci-Guzel, B. Learning to teach argumentation: Case studies of pre-service secondary science teachers. *Eurasia J. Math. Sci.* **2006**, *2* (2), 1−14.

(7) Jiménez-Aleixandre, M. P.; Erduran, S. Argumentation in science education: An overview. In *Argumentation in Science Education*; Springer, 2007; pp 3−27.

(8) Faize, F. A.; Husain, W.; Nisar, F. A critical review of scientific argumentation in science education. *Eurasia J. Math. Sci.* **2017**, *14* (1), 475−483.

(9) Sandoval, W. A.; Enyedy, N.; Redman, E. H.; Xiao, S. Organising a culture of argumentation in elementary science. *Int. J. Sci. Educ.* **2019**, *41* (13), 1848−1869.

(10) Driver, R.; Newton, P.; Osborne, J. Establishing the norms of scientific argumentation in classrooms. *Sci. Educ.* **2000**, *84* (3), 287−312.

(11) de Lima Tavares, M.; Jiménez-Aleixandre, M.-P.; Mortimer, E. F. Articulation of conceptual knowledge and argumentation practices by high school students in evolution problems. *Sci. & Educ.* **2010**, *19* (6), 573−598.

(12) von Aufschnaiter, C.; Erduran, S.; Osborne, J.; Simon, S. Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *J. Res. Sci. Teach.* **2008**, *45* (1), 101−131.

(13) Sevian, H.; Talanquer, V. Rethinking chemistry: a learning progression on chemical thinking. *Chem. Educ. Res. Pract.* **2014**, *15* (1), 10−23.

(14) Talanquer, V. Exploring mechanistic reasoning in chemistry. In *Science Education Research and Practice in Asia-Pacific and Beyond*; Springer, 2018; pp 39−52.

(15) Bricker, L. A.; Bell, P. Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Sci. Educ.* **2008**, *92* (3), 473−498.

(16) Petritis, S. J.; Kelley, C.; Talanquer, V. Exploring the impact of the framing of a laboratory experiment on the nature of student argumentation. *Chem. Educ. Res. Pract.* **2021**, *22* (1), 105−121.

(17) Sadler, T. D. Informal reasoning regarding socioscientific issues: A critical review of research. *J. Res. Sci. Teach.* **2004**, *41* (5), 513−536.

(18) Walker, J. P.; Van Duzor, A. G.; Lower, M. A. Facilitating Argumentation in the Laboratory: The Challenges of Claim Change and Justification by Theory. *J. Chem. Educ.* **2019**, *96* (3), 435−444.

(19) McNeill, K. L.; Krajcik, J. *Book Study Facilitator's Guide: Supporting Grade 5−8 Students in Constructing Explanations in Science: The Claim, Evidence and Reasoning Framework for Talk and Writing*; Pearson Allyn & Bacon: New York, 2012.

(20) Deng, J. M.; Flynn, A. B. Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 749−771.

(21) Cruz-Ramirez de Arellano, D.; Towns, M. H. Students' understanding of alkyl halide reactions in undergraduate organic chemistry. *Chem. Educ. Res. Pract.* **2014**, *15* (4), 501−515.

(22) Lieber, L.; Graulich, N. Investigating Students' Argumentation when Judging the Plausibility of Alternative Reaction Pathways in Organic Chemistry. *Chem. Educ. Res. Pract.* **2022**, *23*, 38−54.

(23) Moon, A.; Stanford, C.; Cole, R.; Towns, M. The nature of students' chemical reasoning employed in scientific argumentation in physical chemistry. *Chem. Educ. Res. Pract.* **2016**, *17* (2), 353−364.

(24) Pabuccu, A.; Erduran, S. Beyond rote learning in organic chemistry: the infusion and impact of argumentation in tertiary education. *Int. J. Sci. Educ.* **2017**, *39* (9), 1154−1172.
(25) Kraft, A.; Strickland, A. M.; Bhattacharyya, G. Reasonable reasoning: multi-variate problem-solving in organic chemistry. *Chem. Educ. Res. Pract.* **2010**, *11* (4), 281−292.
(26) Hosbein, K. N.; Lower, M. A.; Walker, J. P. Tracking Student Argumentation Skills across General Chemistry through Argument-Driven Inquiry Using the Assessment of Scientific Argumentation in the Classroom Observation Protocol. *J. Chem. Educ.* **2021**, *98* (6), 1875−1887.
(27) Walker, J. P.; Sampson, V.; Grooms, J.; Anderson, B.; Zimmerman, C. O. Argument-driven inquiry in undergraduate chemistry labs: The impact on students' conceptual understanding, argument skills, and attitudes toward science. *J. Coll. Sci. Teach.* **2012**, *41* (4), 74−81.
(28) Henderson, J. B.; Osborne, J. F. Using computer technology to support the teaching and learning of argumentation in chemistry. In *Argumentation in Chemistry Education*; Royal Society of Chemistry, 2019; pp 79−105.
(29) Petritis, S. J.; Kelley, C.; Talanquer, V. Analysis of factors that affect the nature and quality of student laboratory argumentation. *Chem. Educ. Res. Pract.* **2022**, *23*, 257−274.
(30) Stowe, R. L.; Cooper, M. M. Arguing from spectroscopic evidence. *J. Chem. Educ.* **2019**, *96* (10), 2072−2085.
(31) Kim, N. J.; Vicentini, C. R.; Belland, B. R. Influence of scaffolding on information literacy and argumentation skills in virtual field trips and problem-based learning for scientific problem solving. *Int. J. Sci. Math. Educ.* **2022**, *20* (2), 215−236.
(32) Wood, D.; Bruner, J. S.; Ross, G. The role of tutoring in problem solving. *J. Child Psychol. Psychiatry* **1976**, *17* (2), 89−100.
(33) Van de Pol, J.; Volman, M.; Beishuizen, J. Scaffolding in teacher−student interaction: A decade of research. *Educ. Psychol.* **2010**, *22* (3), 271−296.
(34) Lieber, L.; Graulich, N. Thinking in Alternatives-A Task Design for Challenging Students' Problem-Solving Approaches in Organic Chemistry. *J. Chem. Educ.* **2020**, *97* (10), 3731−3738.
(35) Lazarou, D.; Erduran, S. Evaluate What I Was Taught, Not What You Expected Me to Know": Evaluating Students' Arguments Based on Science Teachers' Adaptations to Toulmin's Argument Pattern. *J. Sci. Teach. Educ.* **2021**, *32* (3), 306−324.
(36) Toulmin, S. E. *The Uses of Argument*; Cambridge University Press, 2003.
(37) Kang, H.; Thompson, J.; Windschitl, M. Creating Opportunities for Students to Show What They Know: The Role of Scaffolding in Assessment Tasks. *Sci. Educ.* **2014**, *98* (4), 674−704.
(38) Sandoval, W. A.; Millwood, K. A. The quality of students' use of evidence in written scientific explanations. *Cogn. Instr.* **2005**, *23* (1), 23−55.
(39) Choi, A.; Hand, B.; Greenbowe, T. Students' Written Arguments in General Chemistry Laboratory Investigations. *Res. Sci. Educ.* **2013**, *43* (5), 1763−1783.
(40) Walqui, A. Scaffolding instruction for English language learners: A conceptual framework. *Int. J. Biling. Educ. Biling.* **2006**, *9* (2), 159−180.
(41) Kostons, D.; van der Werf, G. The effects of activating prior topic and metacognitive knowledge on text comprehension scores. *Br. J. Educ. Psychol.* **2015**, *85* (3), 264−275.
(42) McNeill, K. L.; Lizotte, D. J.; Krajcik, J.; Marx, R. W. Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *J. Learn. Sci.* **2006**, *15* (2), 153−191.
(43) Talanquer, V. Commonsense chemistry: A model for understanding students' alternative conceptions. *J. Chem. Educ.* **2006**, *83* (5), 811−816.
(44) DeCocq, V.; Bhattacharyya, G. TMI (Too much information)! Effects of given information on organic chemistry students' approaches to solving mechanism tasks. *Chem. Educ. Res. Pract.* **2019**, *20* (1), 213−228.
(45) Linenberger, K. J.; Bretz, S. L. Generating cognitive dissonance in student interviews through multiple representations. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 172−178.
(46) Aimeur, E.; Dufort, H.; Leibu, D.; Frasson, C. Some justifications for the learning by disturbing strategy. In *Artificial Intelligence in Education, 1997: Knowledge and Media in Learning Systems: Proceedings of AI-ED 97, World Conference on Artificial Intelligence in Education, Kobe, Japan*; IOS Press, 1997; Vol. 39, p 119.
(47) Talanquer, V. Concept Inventories: Predicting the Wrong Answer May Boost Performance. *J. Chem. Educ.* **2017**, *94* (12), 1805−1810.
(48) Popova, M.; Bretz, S. L. "It's Only the Major Product That We Care About in Organic Chemistry": An Analysis of Students' Annotations of Reaction Coordinate Diagrams. *J. Chem. Educ.* **2018**, *95* (7), 1086−1093.
(49) Crooks, N. M.; Alibali, M. W. Noticing relevant problem features: Activating prior knowledge affects problem solving by guiding encoding. *Front. Psychol.* **2013**, *4*, 884.
(50) Cohen, J. A power primer. *Psychol. Bull.* **1992**, *112* (1), 155−159.
(51) Lieber, L.; Ibraj, K.; Caspari-Gnann, I.; Graulich, N. Closing the Gap of Organic Chemistry Students' Performance with an Adaptive Scaffold for Argumentation Patterns. *Chem. Educ. Res. Pract.* **2022**, DOI: 10.1039/D2RP00016D.
(52) Luo, X. L.; Wei, B.; Shi, M.; Xiao, X. Exploring the impact of the reasoning flow scaffold (RFS) on students' scientific argumentation: based on the structure of observed learning outcomes (SOLO) taxonomy. *Chem. Educ. Res. Pract.* **2020**, *21* (4), 1083−1094.
(53) Chang, S. N.; Chiu, M. H. Lakatos' scientific research programmes as a framework for analysing informal argumentation about socio-scientific issues. *Int. J. Sci. Educ.* **2008**, *30* (13), 1753−1773.
(54) Windschitl, M. Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Rev. Educ. Res.* **2002**, *72* (2), 131−175.
(55) Caspari-Gnann, I.; Sevian, H. Teacher dilemmas as sources of change and development. *Teach. Teach. Educ.* **2022**, *112*, 103629.
(56) Songer, N. B.; Gotwals, A. W. Guiding explanation construction by children at the entry points of learning progressions. *J. Res. Sci. Teach.* **2012**, *49* (2), 141−165.

https://doi.org/10.1021/acs.jchemed.2c00213
J. Chem. Educ. 2022, 99, 2754−2761

6 Appendices

6.1 Consent Forms and Demographic Questionnaires

6.1.1 Consent Form used at the Justus-Liebig-University Giessen for Study I

Einverständniserklärung zur Auswertung der Interviewergebnisse und Nutzung für wissenschaftliche Publikationen und Konferenzbeiträge

(von Leonie Lieber & Nicole Graulich)

(Oktober/November 2019)

Liebe Studierende,
in diesem Interview möchten wir untersuchen, welche strategischen und kognitiven Ressourcen Sie verwenden und inwiefern diese hilfreich für Sie bei der Bearbeitung der Aufgaben sind. Dabei sollen Rückschlüsse daraus gezogen werden, wie Studierende Aufgaben im Bereich der Organischen Chemie bearbeiten und inwiefern man Studierenden bei der Bearbeitung durch gezielte Aufgaben- und Hilfestellungen unterstützen kann.
Die Teilnahme an diesem Interview hat keine Auswirkungen auf ihre studentischen Leistungen. Das Interview kann jederzeit ohne Angabe von Gründen abgebrochen werden. Alle erhobenen Daten werden im Anschluss an das Interview pseudonymisiert.
Um Ihre Daten anonymisiert auswerten und für wissenschaftliche Zwecke nutzen zu können, füllen Sie bitte folgende Einverständniserklärung aus.

Zutreffendes bitte ankreuzen	**Ich bin damit einverstanden, dass**
☐ Ja ☐ Nein	der von mir ausgefüllte Fragebogen von MitarbeiterInnen der Arbeitsgruppe Graulich ausgewertet werden.
☐ Ja ☐ Nein	die Daten durch studentische Hilfskräfte transkribiert bzw. kodiert werden. Die studentischen Hilfskräfte erfahren dabei zu keinem Zeitpunkt Ihren Namen, sondern nur Ihren Code. Alle beteiligten Personen werden zur Verschwiegenheit gegenüber Dritten verpflichtet.
☐ Ja ☐ Nein	die Daten Gegenstand von wissenschaftlichen Hausarbeiten zum ersten Staatsexamen werden. Die ExamenskandidatInnen erfahren dabei zu keinem Zeitpunkt Ihren Namen, sondern nur Ihren Code. Alle ExamenskandidatInnen werden zur Verschwiegenheit gegenüber Dritten verpflichtet.
☐ Ja ☐ Nein	Screenshots aus den Daten (Transkripte, von Ihnen angefertigte Zeichnungen) in wissenschaftlichen Veröffentlichungen abgebildet werden. Dabei ist zu keinem Zeitpunkt Ihr Name oder Ihr Code dargestellt.

Code in der qualitativen Studie zum Promotionsprojekt von Leonie Lieber (bitte ausfüllen):

Dritter Buchstabe Straße der Adresse
(z. B. Ludwigstraße) ____________

Letzter Buchstabe Vorname einer elterlichen Bezugsperson (vorzugsweise Mutter)
(z. B. Petra oder Peter) ____________

Zweiter Buchstabe Geburtsort
(z. B. Bonn) ____________

Geburtsdatum: Tag (zweistellig)
(z. B. 08.03.1989) ____________

Einwilligungserklärung zur Erhebung und Verarbeitung personenbezogener Daten für Forschungszwecke

1 Gegenstand des Forschungsprojektes

1.	Forschungsprojekt	Untersuchung des Potentials von Aufgaben zur Organischen Chemie im Bezug auf strategische und kognitive Ressourcen
2.	Forschungszweck	Verwendung strategischer und kognitiver Ressourcen bei der Bearbeitung von Aufgaben zur Organischen Chemie im Rahmen der Veranstaltung „Organische Chemie 3 – Katalyse und Synthese“
3.	Durchführende Institutionen	Justus-Liebig-Universität Gießen, Institut für Didaktik der Chemie, Heinrich-Buff Ring 17, 35392 Gießen
4.	Projektleitung	Prof. Dr. Nicole Graulich (JLU Gießen, Tel.: 0641-9934600, Email: nicole.graulich@dc.jlug.de) Leonie Lieber (JLU Gießen, Tel.: 0641-9934611, Email: leonie.lieber@dc.jlug.de)
5.	Erhebungszeitraum	Oktober/November 2019
6.	Interviewerin	Leonie Lieber

2 Einwilligungserklärung

Hiermit willige ich ein, dass im Rahmen des unter (1) beschriebenen Forschungsprojekts Daten meiner Person erhoben und ausgewertet werden. Die Erhebung erfolgt zum einen durch **Video – und Audioaufnahmen**, die in der Folge transkribiert werden. Zum anderen werden **schriftliche Aufzeichnungen** erhoben. Daten aus allen Erhebungsformaten werden vollständig pseudonymisiert und für wissenschaftliche Analysen und in der Folge für wissenschaftliche Veröffentlichungen verwendet und mindestens 10 Jahre lang gespeichert werden.

Über Art und Umfang von Erhebung und Auswertung wurde ich mündlich und in der schriftlichen Anlage zu dieser Erklärung umfassend informiert. Sofern ich Fragen zu dieser vorgesehenen Studie hatte, wurden sie mir vollständig und zu meiner Zufriedenheit beantwortet.

Die Mitglieder der beteiligten Arbeitsgruppen sind zur Verschwiegenheit gegenüber Dritten verpflichtet. Gleichermaßen erkläre ich Verschwiegenheit bezüglich der Aufgabeninhalte gegenüber Dritten.

Die Angaben in der Einverständniserklärung entsprechen meinem freien Willen.

Ihre Einwilligung ist freiwillig. Sie können die Einwilligung ablehnen, ohne dass Ihnen dadurch irgendwelche Nachteile entstehen. Ihre Einwilligung können Sie jederzeit gegenüber der durchführenden Institution widerrufen. Die weitere Verarbeitung Ihrer personenbezogenen Daten wird ab diesem Widerruf unzulässig. Dies berührt jedoch nicht die Rechtmäßigkeit der aufgrund der Einwilligung bis zum Widerruf erfolgten Verarbeitung. Relevante Definitionen der verwendeten datenschutzrechtlichen Begriffe sind in der Anlage Begriffsbestimmungen enthalten.

______________________________	______________	______________________
Vorname und Nachname in Druckschrift	Datum	Unterschrift

6.1.2 Demographic Questionnaire used at the Justus-Liebig-University Giessen for Study I

Demographischer Fragebogen zur statistischen Auswertung der Studienteilnehmenden

(von Leonie Lieber & Nicole Graulich)

(Oktober/November 2019)

Fülle bei den Punkten 1-7 die für dich zutreffenden Angaben aus.

1) Geschlecht ◻ weiblich ◻ männlich ◻ divers

2) Fachsemester ◻ 3 ◻ 5 ◻ 7 ◻ 9 ◻ 11 ◻ anderes: ______

3) Geburtsjahr ____________

4) Ich beherrsche Deutsch auf Muttersprachniveau.

   ◻ ja ◻ nein

5) Ich habe das Modul „Chemie-BK23 Organische Stoffchemie (OC1)“ beim ersten Versuch bestanden.

   ◻ ja ◻ nein

6) Ich habe das Modul „Chemie-BK14 Organische Stoffchemie 2 – Reaktionsmechanismen“ beim ersten Versuch bestanden.

   ◻ ja ◻ nein

7) Ich belege das Modul „Chemie-BV 04 Organische Chemie 3 – Katalyse und Synthese“ zum ersten Mal.

   ◻ ja ◻ nein

6.1.3 Consent Form including a Demographic Questionnaire used at the Tufts University Boston for Study II

Consent-Student-Start Spring 2021

Start of Block: Default Question Block

Q1 **TUFTS UNIVERSITY**
CONSENT TO PARTICIPATE IN A RESEARCH STUDY
Title of the Study: Scaffolding Argumentation Patterns of Organic Chemistry Students
Principal Investigator: Assistant Professor Ira Caspari, Tufts University, Department of Chemistry
Phone: 623-330-0687
Email: ira.caspari@tufts.edu

You are being asked to volunteer in a research study. Please find below information about this research for you to carefully consider when deciding about whether or not to participate. Please ask questions about any of the information you do not understand before you decide whether to participate.

What is this study about?
Researchers at Tufts University (Dr. Caspari) and the Justus-Liebig-University Giessen, Germany (Leonie Lieber and Dr. Nicole Graulich) are conducting a study on organic chemistry students' argumentation patterns. The purpose of the research is to better understand how students construct arguments on their own and with support by a scaffold. Your participation in this study will help to investigate how organic chemistry students can be supported in their argumentation patterns. You are asked to participate because you are a student in the Chem 52 course in which the research will take place. You are one of about 70 participants to take part in this study. You will be enrolled in this study for one month during which you will be participating in two surveys two weeks apart.

What will happen during this research?
If you agree to be in this research, your participation will include:

- Answer two surveys two weeks apart. Answering each survey will last about 90 minutes. In the surveys, you will be shown organic reactions and you will be asked to make predictions about those reactions. The surveys will support you in constructing arguments in order to support your reasoning about the reactions. The first survey will be the same for every student who participates in the study. The second survey will be individualized and support you in the area that you might benefit the most.

- Answer some demographics questions at the end of this consent form. You can opt not to answer the demographics questions and still participate in the study.

Page 1 of 5

What will you do to protect my privacy?
Before sharing any data outside the research team, we will de-identify it so that the only way it can be directly identified is with a key to which only Dr. Caspari has access. Only Dr. Caspari will have access to data in which participants are identifiable. We will store digital data on secure password protected research drives or in Tufts Box. Transcripts of your survey answers and drawings you will upload within the survey may be reproduced in whole or in part for use in presentations, written products, or professional development that result from this study. Only the demographics information that you choose to disclose in the demographics survey below will be used in presentations or written products resulting from this study. Neither your name nor any identifying information that you do not disclose in the survey will be used in presentations or in written products resulting from the study.

We will handle your data carefully, but we can't guarantee that your privacy will be protected. Individuals and organizations, including the Tufts Social, Behavioral & Educational Research Institutional Review Board (SBER-IRB), which are responsible for conducting or monitoring this research, may be permitted access to and inspect the research records. If you tell us something that makes us believe that you or others have been or may be physically harmed, we may report that information to the appropriate agencies. Also, your de-identified data could be used for future research studies or distributed to another investigator for future research studies without your additional informed consent.

What are the risks or discomforts associated with this research?
You might feel uncomfortable during the survey because you are sharing your science thinking. The risk of feeling uncomfortable is minimal as you can choose to not answer any given questions and/or stop your participation at any time.

How might I benefit from this research?
Participating in these surveys may support you to improve your argumentation skills in organic chemistry.

What is the compensation for the research?
You will not receive any compensation for your participation in study.

What will happen if I choose not to participate?
It is your choice to participate or not to participate in this research. Participation is voluntary. If you decide not to participate, you do not have to do the survey.

Is my participation voluntary, and can I withdraw?
Your participation in this study is voluntary, and you can withdraw at any time by contacting the PI of the study. If you withdraw, the data collected up to that point will still be used unless you specify otherwise.

Who do I talk to if I have questions?
If you have questions or concerns, contact the research team at:

Ira Caspari
623-330-0687
ira.caspari@tufts.edu

An Institutional Review Board ("IRB") is overseeing this research. An IRB is a group of people who perform independent review of research studies to ensure the rights and welfare of participants are protected. If you have questions about your rights or wish to speak with someone other than the research team, you may contact:
Tufts Social, Behavioral, and Educational Research IRB
75 Kneeland Street, Suite 623
Boston, MA 02111
617.627.8804
SBER@tufts.edu

STATEMENT OF CONSENT I have read and considered the information presented in this form. I confirm that I understand the purpose of the research and the study procedures. I understand that I may ask questions at any time and can withdraw my participation without prejudice. I have read this consent form.

◯ I consent to participate in this study.

Q2 Please write your full name here.

____

Q3 Please write your e-mail address here.

____

User Code

Third letter of the street of your adress (e.g. 20 In**G**ram Street, Queens)

Last letter of the first name of a parent/guardian (if you have two, choose the older one) (e.g., Mar**Y** or Be**N**)

Second letter of the city/town/village your parent/guardian lives in (if you have two, choose the older one) (e.g., M**E**dford)

Date of birth: Day (two digits) (e.g., 08)

Q4
Demographic questions

The following questions are intended to learn about your demographics and will be used in conjunction with the main data that will be collected for this study. It will take you no more than 5 minutes to answer all questions. For each question, you can also choose not to answer.

Q5 Please indicate your gender.

- ○ Non-binary
- ○ Female
- ○ Male
- ○ Other/Prefer to self-describe (please list below)
  ____
- ○ Prefer not to answer

Q6 Please select the range of ages in which you fall.

- ○ <18
- ○ 18-22
- ○ 23-26
- ○ 27-30
- ○ >30
- ○ Prefer not to answer

Q7 With which ethnicities do you identify? Please use your own understanding of the below identifiers. (You may choose multiple and/or self-identify.)

- ☐ Native American/Alaskan Native
- ☐ Asian
- ☐ Black
- ☐ Latino/Latinx
- ☐ Pacific Islander
- ☐ White
- ☐ Other/Prefer to self-describe (please list below)

____

- ☐ Prefer not to answer

Q8 Do you identify as Hispanic?

- ○ Yes
- ○ No
- ○ Prefer not to answer

Q9 Was this class taught in the language you are most comfortable with?

- ○ No (I feel most comfortable speaking/learning in another language)
- ○ Yes (I feel most comfortable speaking/learning in this language)
- ○ I would feel equally comfortable speaking/learning in this language or another
- ○ Prefer not to answer

6.2 Supporting Information Study I

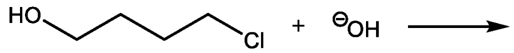
6.2.1 Original Tasks from the Qualitative Interview Study

1) Welches Produkt wird bei der Reaktion gebildet?

Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

2) Welches Produkt wird bei der Reaktion gebildet?



Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

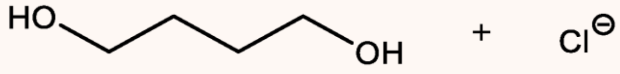
1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

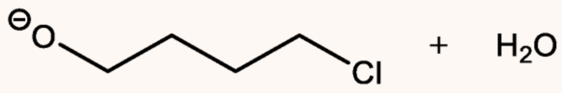
3) Beschreibe, wie du denkst, dass die Studierenden auf diese Lösung gekommen sind. Was glaubst du war ihr Gedankengang?

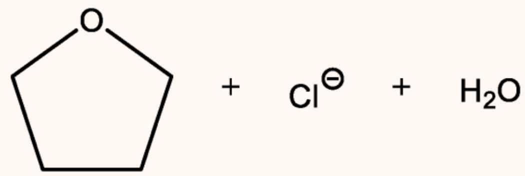
Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

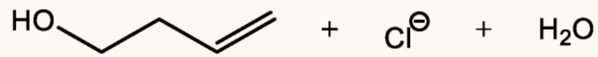
1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

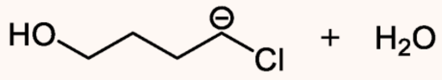
4) Bleiben Sie bei Ihrem Produkt aus 2) oder entscheiden Sie sich für eine/mehrere der anderen Lösungen?

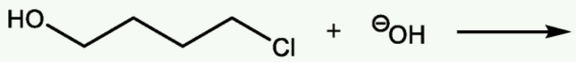

HO
OH
+
Cl⊖

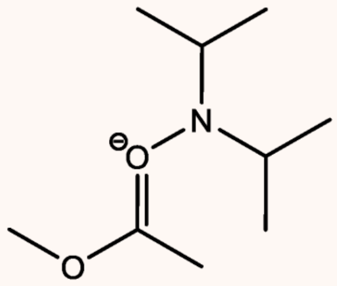

⊖O
Cl
+
H2O

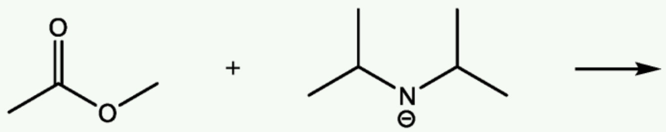

O
+
Cl⊖
+
H2O

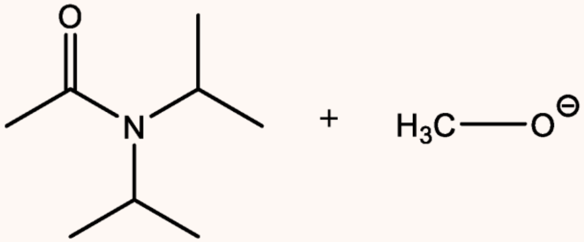

HO
+
Cl⊖
+
H2O

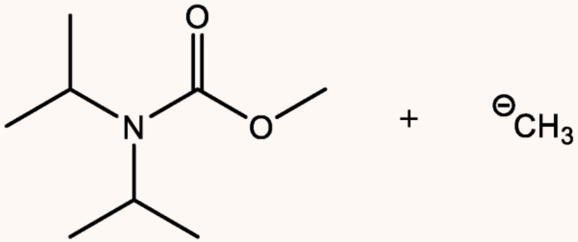

HO
Cl
+
H2O

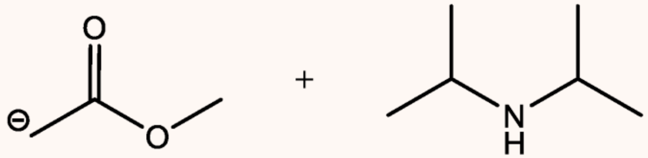

HO
Cl
+
OH

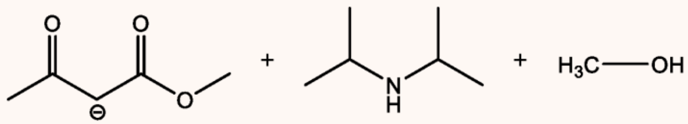

N
O
O
O


O
O
+
N

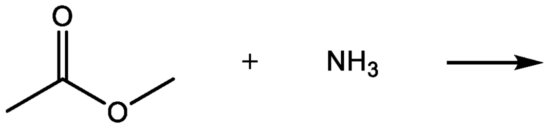








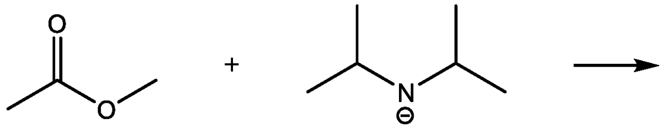
1) Welches Produkt wird bei der Reaktion gebildet?



Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

2) Welches Produkt wird bei der Reaktion gebildet?



Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

3) Beschreibe, wie du denkst, dass die Studierenden auf diese Lösung gekommen sind. Was glaubst du war ihr Gedankengang?

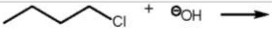
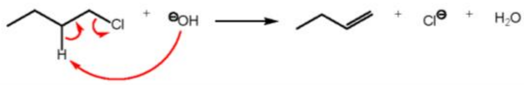
Wie würdest du die Schwierigkeit der Aufgabe anhand deiner Bearbeitung auf einer Skala von 1 – 5 bewerten? (1 = leicht, 5 = schwer)

1 (leicht)	2 (eher leicht)	3 (mittel)	4 (eher schwer)	5 (schwer)
☐	☐	☐	☐	☐

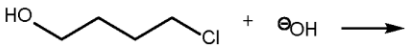
4) Bleiben Sie bei Ihrem Produkt aus 2) oder entscheiden Sie sich für eine/mehrere der anderen Lösungen?

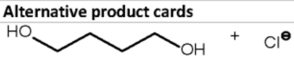
6.2.2 Sample Solutions of the Tasks from the Qualitative Interview Study

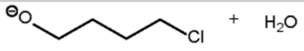
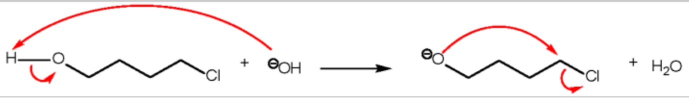
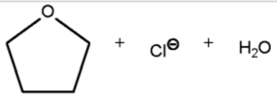
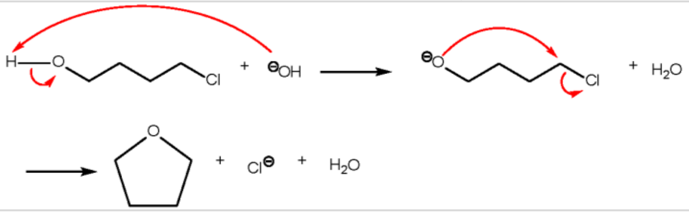
Source of ideas regarding the task: Vollhardt (2011), p. 262

	Option 1:  Option 2: 
	- Option 1: Nucleophilic substitution (S_N2), because of a primary C-atom (because the carbocation is not stabilized), Cl^- is a good leaving group (because Cl^- is a weak base and has a good ability to pick up a negative charge), the most electrophilic site of the alkane is at the α-C-atom, OH^- is a strong nucleophile (because it has a negative charge and a low delocalization of this negative charge) and thus, reacts at the α-C-atom (OH^- reacts at the C-atom on the opposite site of the leaving group, because the antibonding σ*-orbital is on the opposite side as the C-Cl bond, which has a big orbital lobe, the reaction of the nucleophile on the opposite site of the C-atom is preferred), thereby the chloride ion is split off concertedly (formation of a trigonal-bipyramidal transition state – energetic maximum) and the bonding electron pair of the carbon atom moves to the leaving group. - When the nucleophile approximates the sp^3 – hybridized center and the transition state is formed the molecule becomes planar (sp^2), with the chloride ion leaving the molecule the tetrahedral sp^3- hybridization is formed again. - Option 2: β- elimination (E2), OH^- serves as a base and deprotonates the β-C-atom (the hydrogen is acidic, because the anion is stabilized by the overlap with the anti-bonding C-Cl-orbital), whereby the chloride ion is split off concertedly and a double bond is formed. - Conformation must be synperiplanar or antiperiplanar, since these are the only conformations that allow a coplanar transition state of the σ-electrons of the substrate into the π-electrons of the alkene (antiperiplanar generates an energetically favorable state since the staggered transition state), the re-hybridization of both carbon centers from sp^3 to sp^2 generates two p-orbitals available for forming a double bond.

Information for use: The following alternative product cards are neither a complete illustration of possible products nor necessarily correct but represent typical students' solutions. The explanations refer back to the predicted products and mechanisms. For use in class, the explanations can be extended or reduced as needed.

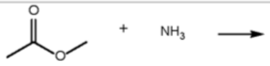
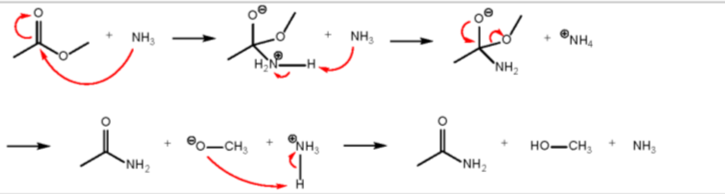


Alternative product cards	Mechanisms for the product cards + explanations for the mechanisms
	
	- S_N2-reaction analogous to part A1 of the task - Nucleophilic substitution (S_N2), because of a primary C-atom (because the carbocation is not stabilized), Cl^- is a good leaving group (because Cl^- is a weak base and has a good ability to pick up a negative charge), the most electrophilic site of the alkane is at the α-C-atom, OH^- is a strong nucleophile (because it has a negative charge and a low delocalization of this negative charge) and thus, reacts at the α-C-atom (OH^- reacts at the C-atom on the opposite site of the leaving group, because the antibonding σ*-orbital is on the opposite side as the C-Cl bond, which has a big orbital lobe, the reaction of the nucleophile on the opposite site of the C-atom is preferred), thereby the chloride ion is split off concertedly (formation of a trigonal-bipyramidal transition state – energetic maximum) and the bonding electron pair of the carbon atom moves to the leaving group. - When the nucleophile approximates the sp^3 – hybridized center and the transition state is formed the molecule becomes planar (sp^2), with the chloride ion leaving the molecule the tetrahedral sp^3-hybridization is formed again. - OH^- was only considered as a nucleophile not as a base. However, the most acidic hydrogen is located at the hydroxyl group of the alkane. Acid-base reactions proceed faster than substitutions.

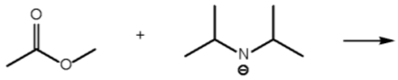
	 - The acid-base reaction between alkoxide as the base and water as the acid proceeds faster than substitutions or eliminations because there is no need for a change of geometry during an acid-base reaction. Protons are small atoms therefore they transport little mass which is why the activation energy is very low (however, equilibrium hence the pK_a, alcohol and the water have similar pK_a) - The creation of the alkoxide ion leads to the formation of another nucleophile (pK_b of – 1,9, OH^- pK_b of - 1,7) which can react at the molecule on the most electrophilic site (at the α-C-atom)
	 - The acid-base reaction between alkoxide as the base and water as the acid proceeds faster than substitutions or eliminations because there is no need for a change of geometry during an acid-base reaction. Protons are small atoms therefore they transport little mass which is why the activation energy is very low (however, equilibrium hence the pK_a, alcohol and the water have similar pK_a) - The creation of the alkoxide ion leads to the formation of another nucleophile (pK_b of – 1,9, OH^- pK_b of - 1,7) which can react at the molecule on the most electrophilic site (on the α-C-atom) - Nucleophilic substitution (S_N2), because of primary C-atom (because the carbocation is not stabilized), Cl^- is a good leaving group (because Cl- is a weak base and has a good ability to pick up a negative charge), the most electrophilic site of the alkane is at the α-C-atom, $R\text{-}O^-$ is a strong nucleophile (because it has a negative charge and a low delocalization of this negative charge), $R\text{-}O^-$ reacts at the α-C-atom (OH^- reacts at the C-atom on the opposite site of the leaving group, because the antibonding σ*-orbital is on the opposite side as the C-Cl bond which has a big orbital lobe, the reaction of the nucleophile on the opposite site of the C-atom is preferred), thereby the chloride ion is split off concertedly (formation of a

References:

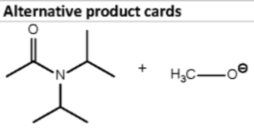
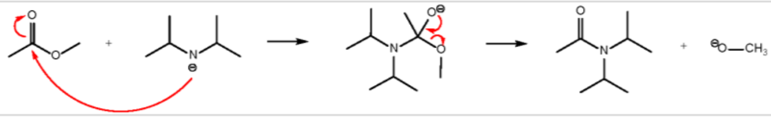
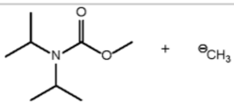
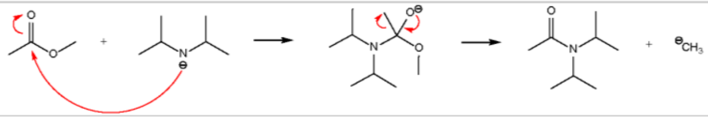
1. Buddrus, J. S., Bernd, *Grundlagen der organischen Chemie*. 2015; Vol. 5.
2. Vollhardt, K. P. C. S., Neil. E, *Organische Chemie*. 2011; Vol. 5.
3. Carey; F.A.; Sundberg, R. J., *Organische Chemie*. WILEY-VCH: Weinheim, 1995.
4. Heine, H. W.; Miller, A. D.; Barton, W. H.; Greiner, R. W., On Cyclic Intermediates in Substitution Reactions. IV. The Hydrolysis of Trimethylene-and Tetramethylene Chlorohydrins. *Journal of the American Chemical Society* **1953,** *75* (19), 4778-4779.

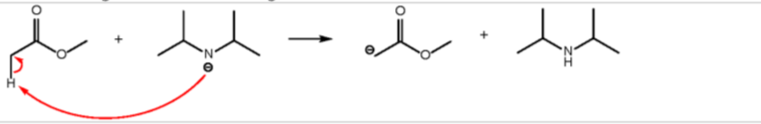
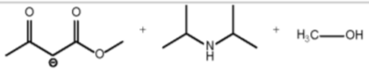
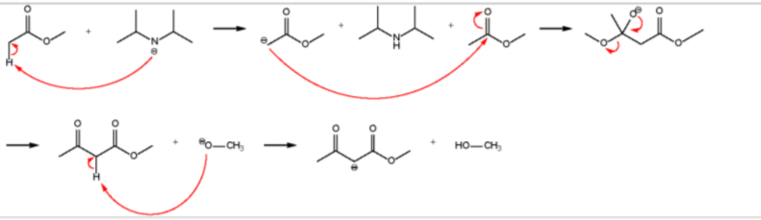
	
	- Ammonia is strongly nucleophilic though it is a rather weak base, while reacting at the electrophilic carbonyl-C-atom a change in hybridization occurs (sp^2 to sp^3, change of geometry from 120° to 109,5°) and the electron pair of the C=O- double bond is moved to the oxygen atom (π-bond is broken), a σ-bond is formed between the carbonyl-C-atom and the nucleophile; a tetrahedral intermediate is created. In this case ammonia is a better leaving group (pK_a of the conjugate acid: NH_4^+ 9,25, methanol 15). Due to deprotonation of the amine molecule by a second ammonia molecule in basic conditions the methoxide ion becomes the better leaving group (pK_a of the conjugate acid: NH_3 23, methanol 15). The lone pair of the nitrogen atom supports the bond breaking of the methoxide-C-bond. Through the renewed formation of the C=O- double bond the methoxide ion is split off. An amide group is created in the product.

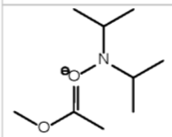
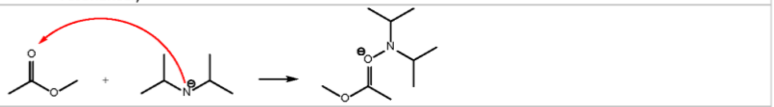
Information for use: The following alternative product cards are neither a complete illustration of possible products nor necessarily correct but represent typical students' solutions. The explanations refer back to the predicted products and mechanisms. For use in class, the explanations can be extended or reduced as needed.



LDA = lithium diisopropylamide

Alternative product cards	Mechanisms for the product cards + explanations for the mechanisms
	
	- LDA reacts at the electrophilic carbonyl-C-atom and the electron pair of the C=O-double bond is moved to the oxygen atom (π-bond is broken), a σ-bond between the carbonyl-C-atom and the nucleophile is formed; a tetrahedral intermediate is created, the carbonyl-C-atom re-hybridizes from sp^2 to sp^3 (change of geometry from 120° to 109,5°). Through the renewed formation of the C=O- double bond the methoxide ion is split off. An amide group is created in the product. - This scenario is implausible, hence LDA is sterically hindered und therefore has not enough space to react at the carbonyl-C atom (Bürgi-Dunitz-Angle 107° too small for LDA), acid-base reaction is faster than a substitution
	
	- LDA reacts at the electrophilic carbonyl-C-atom and the electron pair of the C=O-double bond is moved to the oxygen atom (π-bond is broken), a σ-bond between the carbonyl-C-atom and the nucleophile is formed; a tetrahedral intermediate is created, the carbonyl-C-atom re-hybridizes from sp^2 to sp^3

	(change of geometry from 120° to 109,5°). Through the renewed formation of the C=O- double bond the carbanion ion is split off. - This scenario is implausible, hence LDA is sterically hindered und therefore has not enough space to react at the carbonyl-C atom (Bürgi-Dunitz-Angle 107° too small for LDA). Additionally, carbanion is a very bad leaving group since CH_3^- is a strong base (pK_a of CH_4 = 56) and does not easily accommodate the negative and localized charge.
	 - LDA is a strong base and a sterically hindered nucleophile, hence the required space of the nucleophile hinders a successful reaction at the electrophilic carbonyl-C-atom. - In an acid-base reaction LDA, serving as a base, reacts at the most acidic site of the ester molecule (hydrogen on the α-C-atom) and deprotonates it. ○ With the protonation LDA creates a diisopropylamine molecule - Through the deprotonation of the α-carbon atom an ester-enolate can be formed which is stabilized by resonance.
	 - LDA is a strong base and a sterically hindered nucleophile, hence the required space of the nucleophile hinders a successful reaction at the electrophilic carbonyl-C-atom. - In an acid-base reaction LDA, serving as a base, reacts at the most acidic site of the ester molecule (hydrogen on the α-C-atom) and deprotonates it.

	○ With the protonation LDA creates a diisopropylamine molecule - Through the deprotonation of the α-carbon atom an ester-enolate can be formed which stabilized by resonance. The enolate molecule is a strong nucleophile because it has a negative charge compared to the negative partial charge of the other nucleophilic centers (pK_a alkane approximately 50, alcohol approximately 15). Thus, the enolate reacts at a second ester molecule on the most electrophilic site (carbonyl-C atom), the electron pair of the C=O- double bond is moved to the oxygen atom (π-bond is broken), a σ-bond is formed between the carbonyl-C and the nucleophile, thereby a tetrahedral intermediate is created, the carbonyl-C-atom re-hybridizes from sp^2 to sp^3 (change of geometry from 120° to 109,5°). Through the renewed formation of the C=O- double bond the methoxide ion is split off. Methoxide is a strong base and during an acid-base reaction deprotonates the most acidic site (hydrogen at the α-C-atom in between the carbonyl groups, hence the anion is stabilized by resonance.)
	 - Implausible since the most nucleophilic site of the LDA reacts at the most nucleophilic site of the ester molecule, the oxygen molecule does not obey the octet rule (10 e^-, four orbitals only that would be completely filled with 8 electrons)

References:

1. Buddrus, J.S., Bernd, *Grundlagen der organischen Chemie*. Vol. 5. 2015.
2. Vollhardt, K.P.C.S., Neil. E, *Organische Chemie*. Vol. 5. 2011.
3. Carey; F.A.; Sundberg, R.J., *Organische Chemie*. 1995, Weinheim: WILEY-VCH.
4. McClelland, R.A., *Kinetics and mechanism of amide acetal hydrolysis. Carbon-oxygen vs. carbon-nitrogen bond cleavage in acid solutions.* Journal of the American Chemical Society, 1978. **100**(6): p. 1844-1849.

6.2.3 Interview Protocol

Interviewprotokoll

Vorab

- Datenschutzerklärung ausfüllen und weglegen
- Einverständniserklärung erläutern (die Daten werden voll anonymisiert und sind nicht nachvollziehbar)
    - Sarah wird zu A4VG29 wird zu Anna
    - Man erkennt sich bestenfalls selbst nicht wieder
- Code generieren
    - Elterliche Bezugsperson: vorzugsweise Mutter
- Demographischer Fragebogen
- Einweisung in lautes Denken
    - Es geht nicht darum, dass Aufgaben richtig oder falsch gelöst werden
    - Es ist wichtig zu erfahren welche Gedankengänge Studis bei der Bearbeitung der Aufgaben haben, was ihnen leichtfällt oder schwer
    - Es ist wichtig, ehrlich und offen zu sein, das ist keine Bewertungssituation

Aufgabenteil 1+2

- **Wie bist du auf dieses Produkt gekommen?**
- Was war deine Vorgehensweise?
    - **Warum hast du das so gemacht?**
        - **IMPLIZIT vs EXPLIZIT**
- Was ist dir durch den Kopf gegangen?
- Kannst du mir anhand einer Darstellung zeigen, was du gerade mit … gemeint hast?
    - **Kannst du das noch mal aufzeichnen?**
- Du hast gerade … gesagt. Was meinst du damit in diesem Kontext?
- Ich weiß, du hast das schon einmal erläutert, aber du hilfst mir das besser zu verstehen, wenn du das jetzt noch mal beschreibst.

Aufgabenteil 3

- Die Karten sind nur eine Auswahl an möglichen Produkten und sind nicht vollständig
- **Welche Aspekte erachtest du als plausibel und welche als nicht plausibel? Warum?**
- Was war deine Vorgehensweise?
    - **Warum hast du das so gemacht?**
        - **IMPLIZIT vs EXPLIZIT**
- Was ist dir durch den Kopf gegangen?
- Kannst du mir anhand einer Darstellung zeigen, was du gerade mit … gemeint hast?
    - **Kannst du das noch mal aufzeichnen?**
- Du hast gerade … gesagt. Was meinst du damit in diesem Kontext?
- Ich weiß, du hast das schon einmal erläutert, aber du hilfst mir das besser zu verstehen, wenn du das jetzt noch mal beschreibst.

Aufgabenteil 4

- **Warum hast du dich für dieses/diese Produkte entschieden?**
- Was war deine Vorgehensweise?
    - Warum hast du das so gemacht?

- IMPLIZIT vs EXPLIZIT

- Was ist dir durch den Kopf gegangen?
- Du hast gerade … gesagt. Was meinst du damit in diesem Kontext?
- Ich weiß, du hast das schon einmal erläutert, aber du hilfst mir das besser zu verstehen, wenn du das jetzt noch mal beschreibst.
- **Sortiere dein Produkt und alle anderen gegebenen Produkte bitte nach Plausibilität**

Evaluation

- Likertskala beantworten
- Wie waren die Aufgaben für dich?
- Was ist deine Meinung zu diesem Aufgabentyp?
- **Hattest du das Gefühl, dass du zum Beispiel chemische Konzepte eher überdacht hast als in traditionellen Aufgaben?**
- Haben dir die Aufgaben geholfen?

6.2.4 Students' Argumentation Patterns

The following argumentation patterns derived from the think-aloud interviews for study I in October and November 2019. The argumentation patterns were created for each participant and illustrate their argumentation when judging the plausibility of alternative reaction pathways. Thereby, students' utterances show the experienced argumentation process and does not guarantee technical correctness.

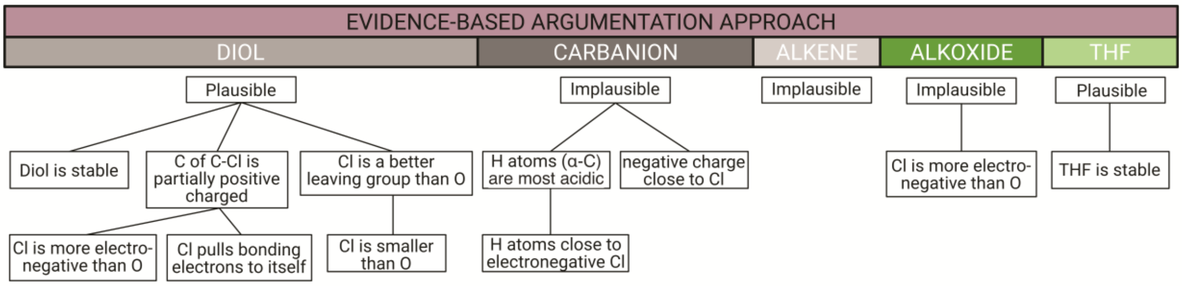


Fig. 13: Amber's argumentation pattern of the evidence-based argumentation approach.

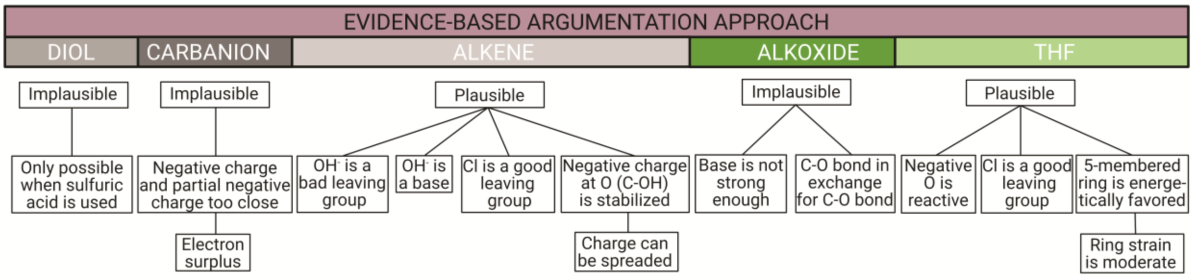


Fig. 14: Haley's argumentation pattern of the evidence-based argumentation approach.

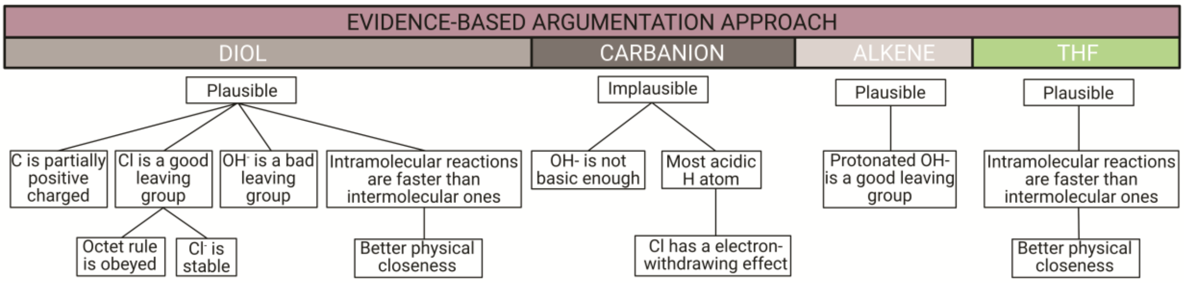


Fig. 15: Gloria's argumentation pattern of the evidence-based argumentation approach.

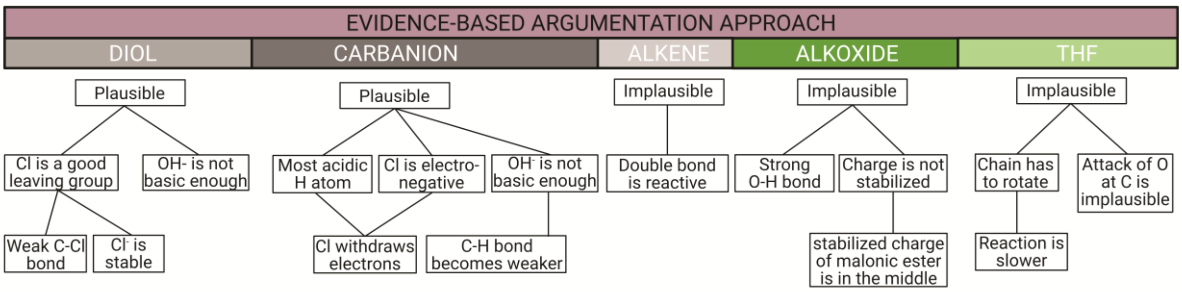


Fig. 16: Pam's argumentation pattern of the evidence-based argumentation approach.

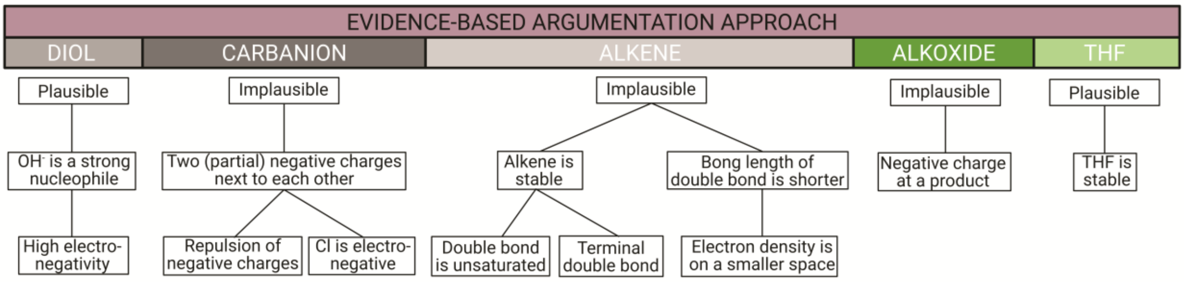


Fig. 17: Sonia's argumentation pattern of the evidence-based argumentation approach.

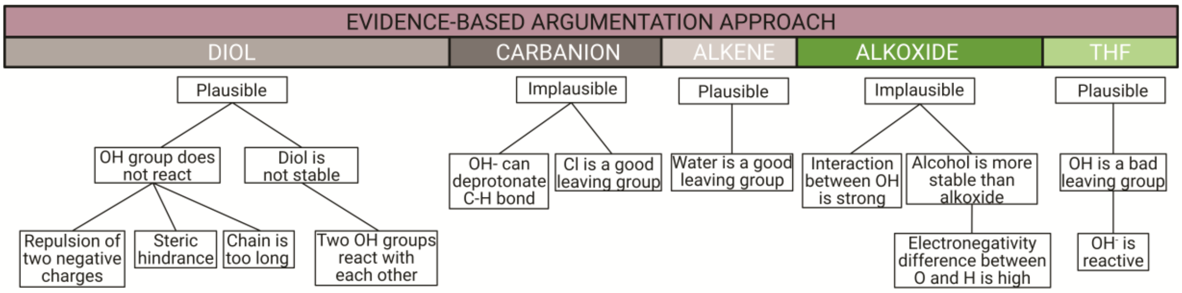


Fig. 18: Sal's argumentation pattern of the evidence-based argumentation approach.

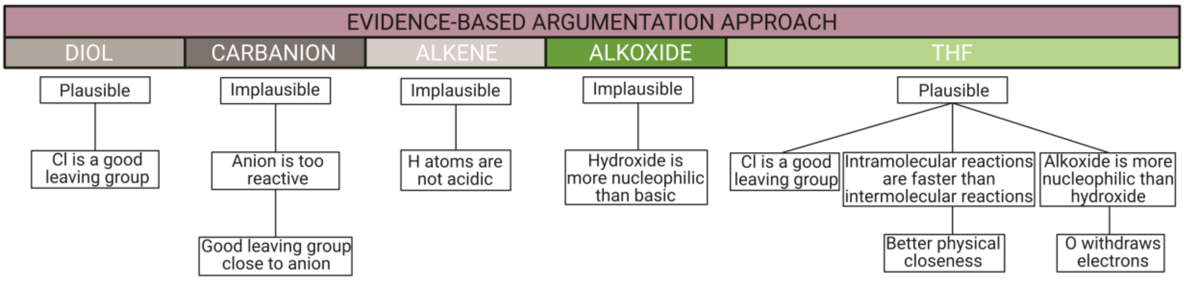


Fig. 19: Frank's argumentation pattern of the evidence-based argumentation approach.

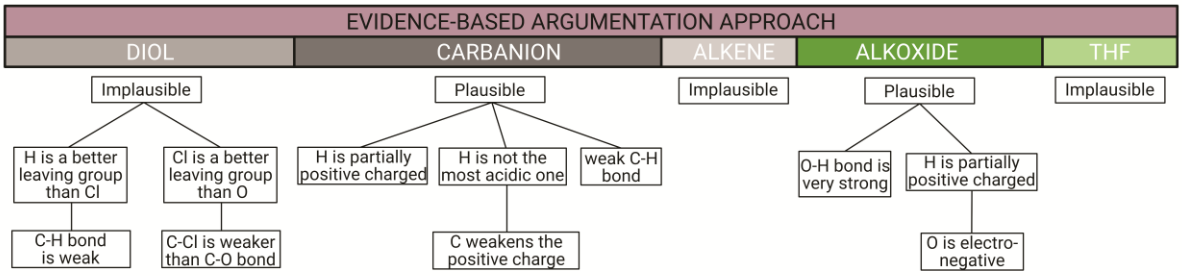


Fig. 20: Beth's argumentation pattern of the evidence-based argumentation approach.

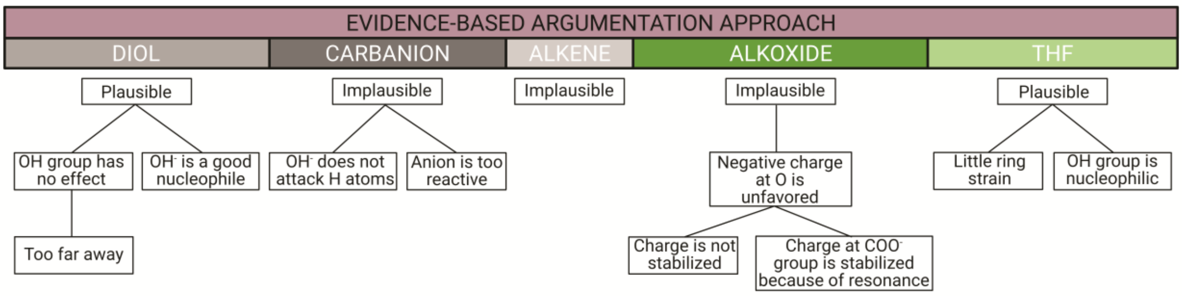


Fig. 21: Andy's argumentation pattern of the evidence-based argumentation approach.

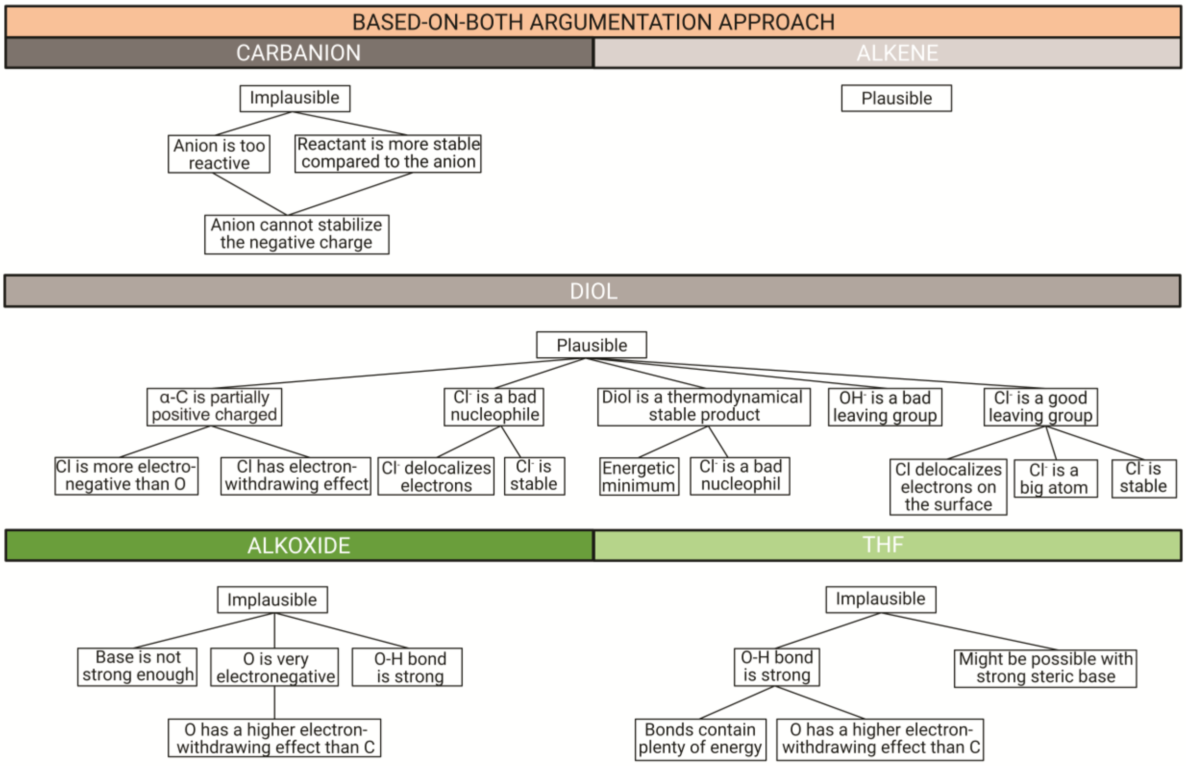


Fig. 22: Ben’s argumentation pattern of the based-on-both argumentation approach.

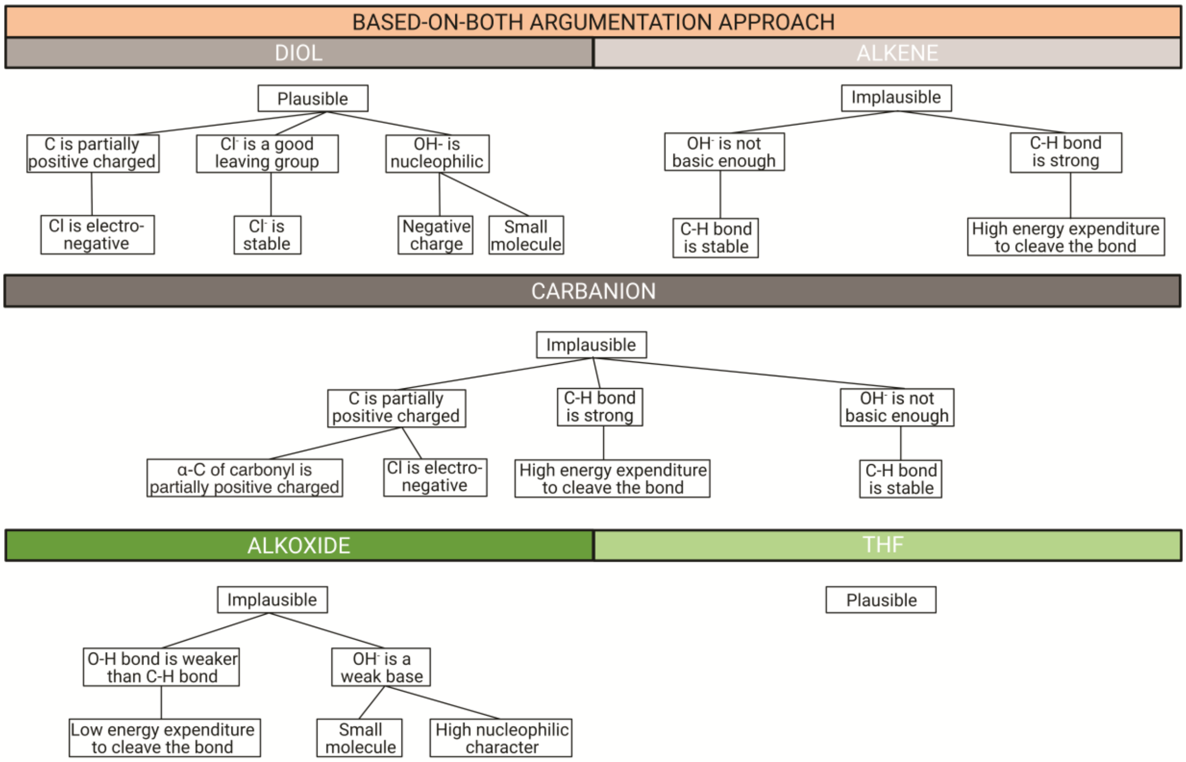


Fig. 23: Joe’s argumentation pattern of the based-on-both argumentation approach.

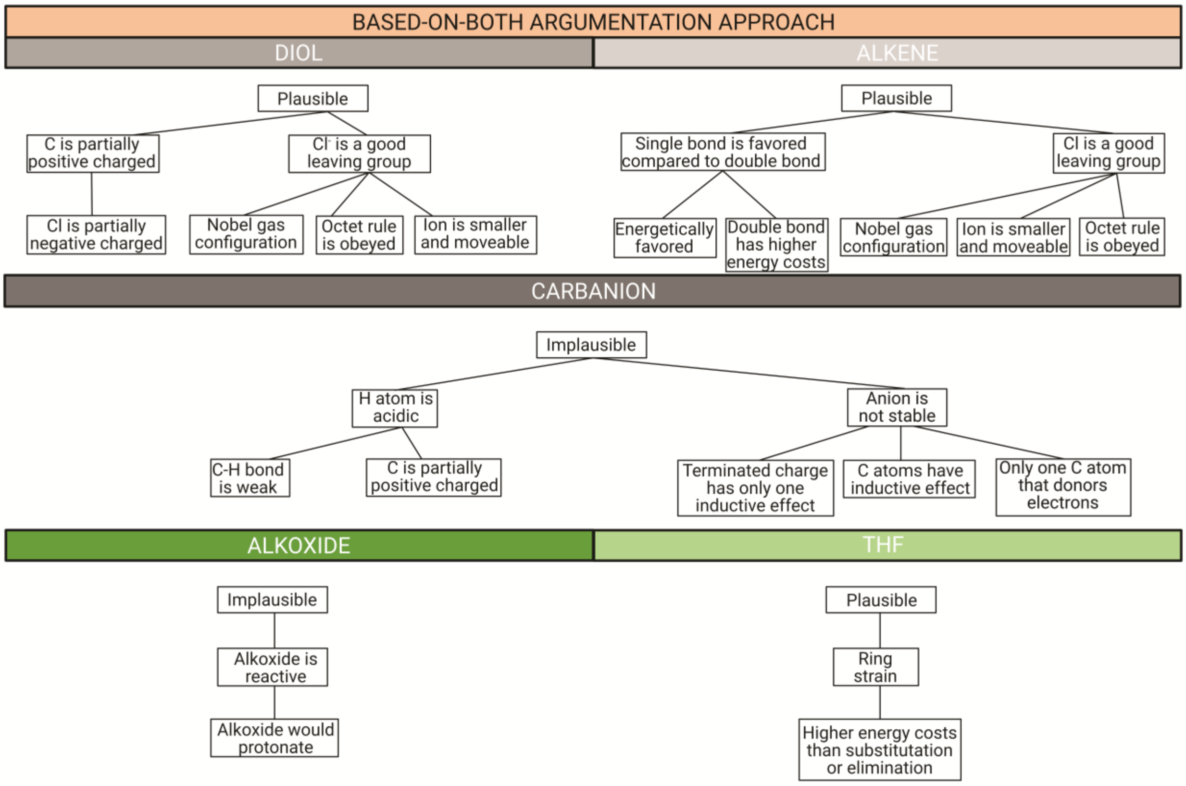


Fig. 24: Stella's argumentation pattern of the based-on-both argumentation approach.

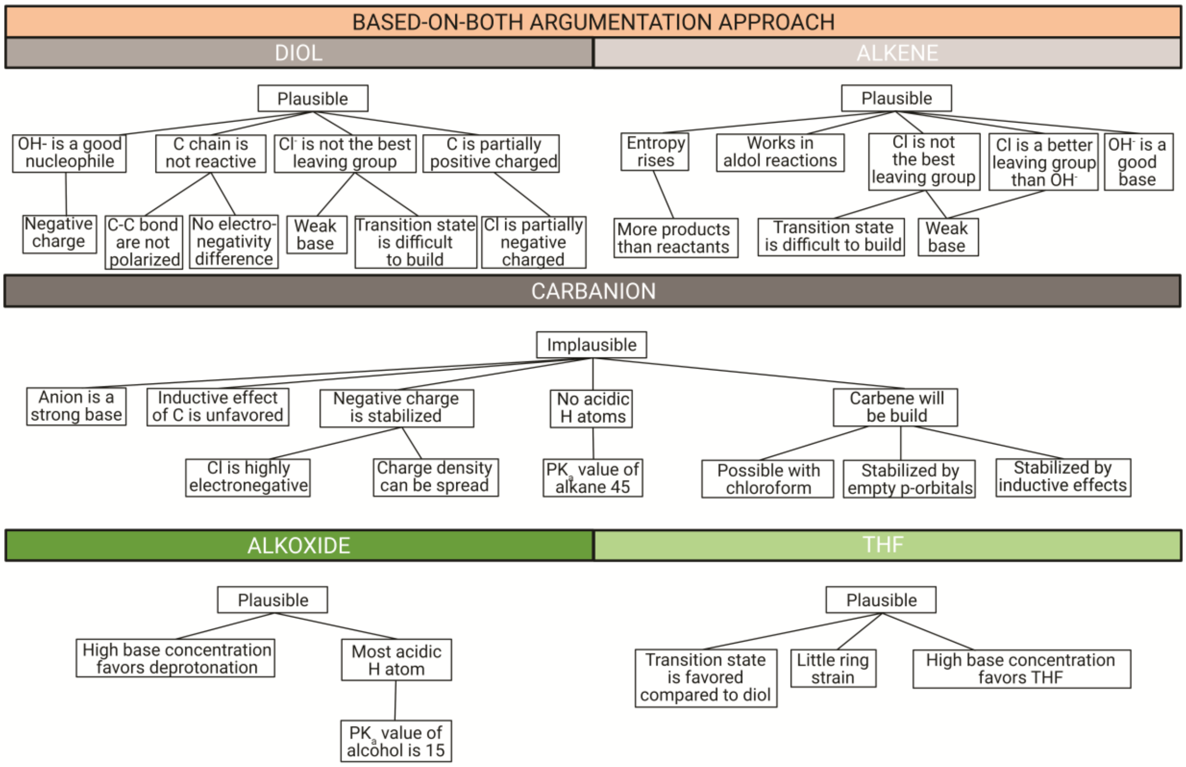


Fig. 25: Manny's argumentation pattern of the based-on-both argumentation approach.

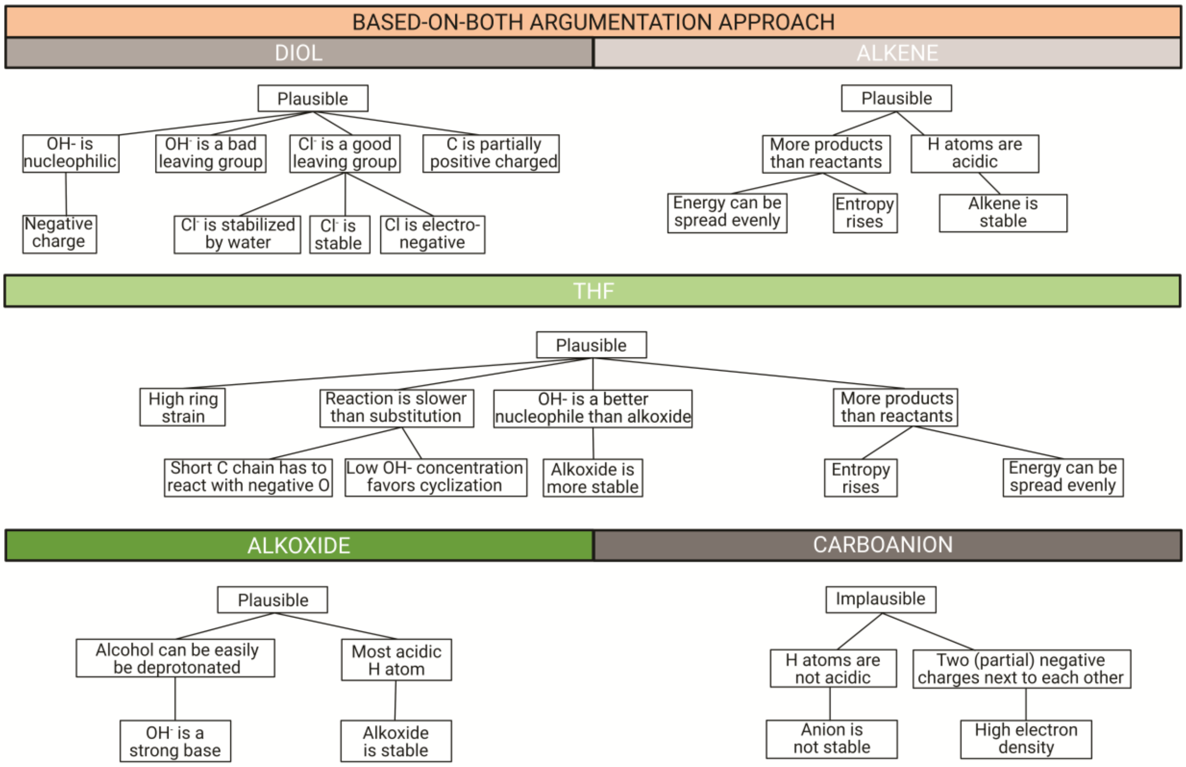


Fig. 26: Mitchell's argumentation pattern of the based-on-both argumentation approach.

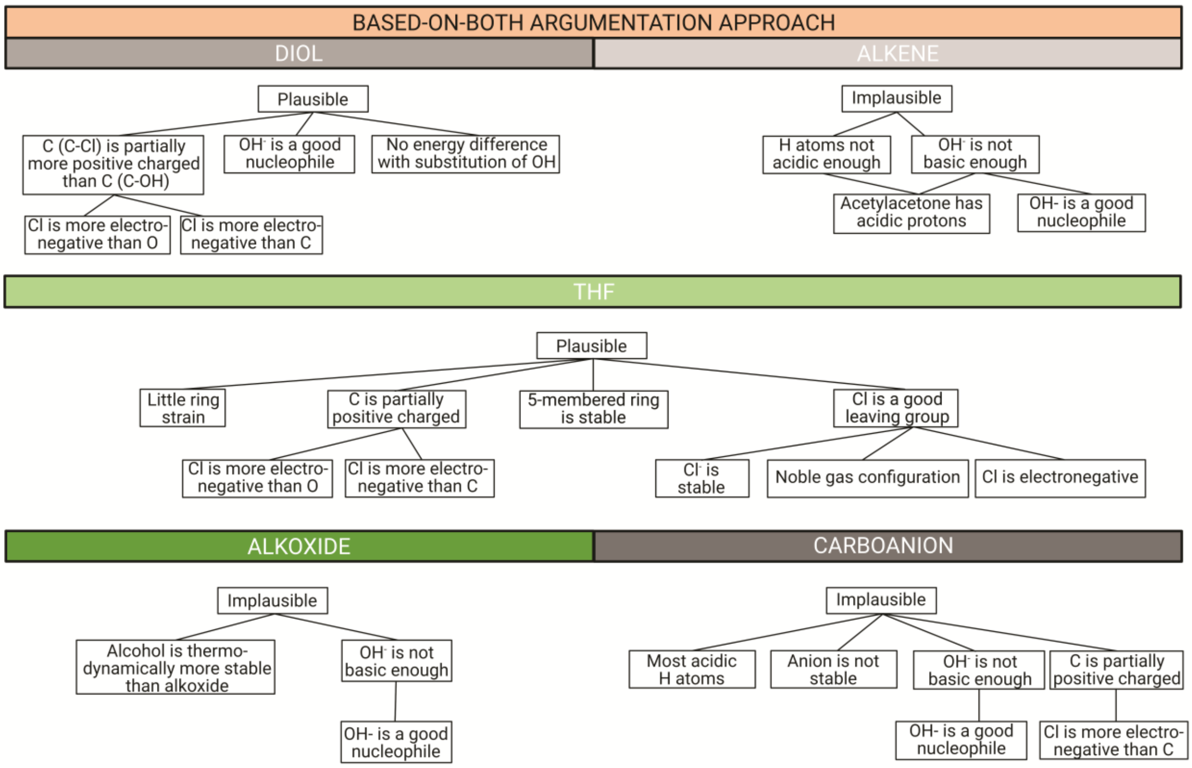


Fig. 27: Dylan's argumentation pattern of the based-on-both argumentation approach.

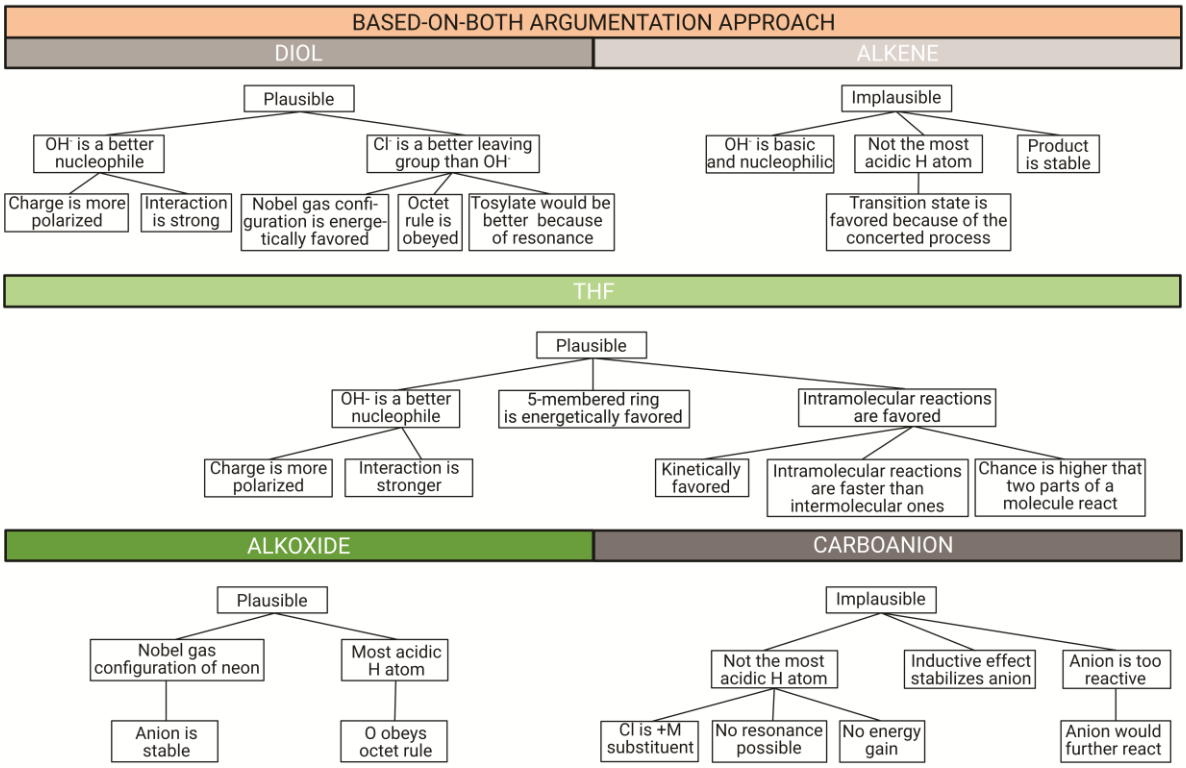


Fig. 28: Gil's argumentation pattern of the based-on-both argumentation approach.

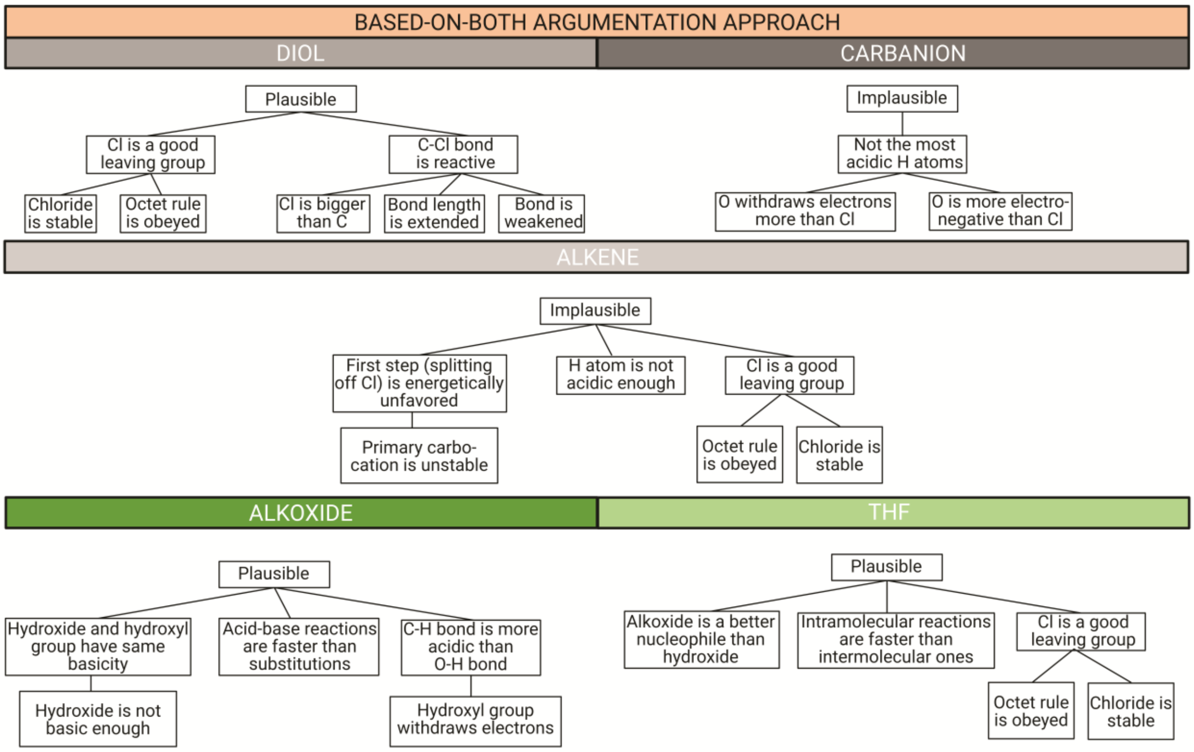


Fig. 29: Sydney's argumentation pattern of the based-on-both argumentation approach.

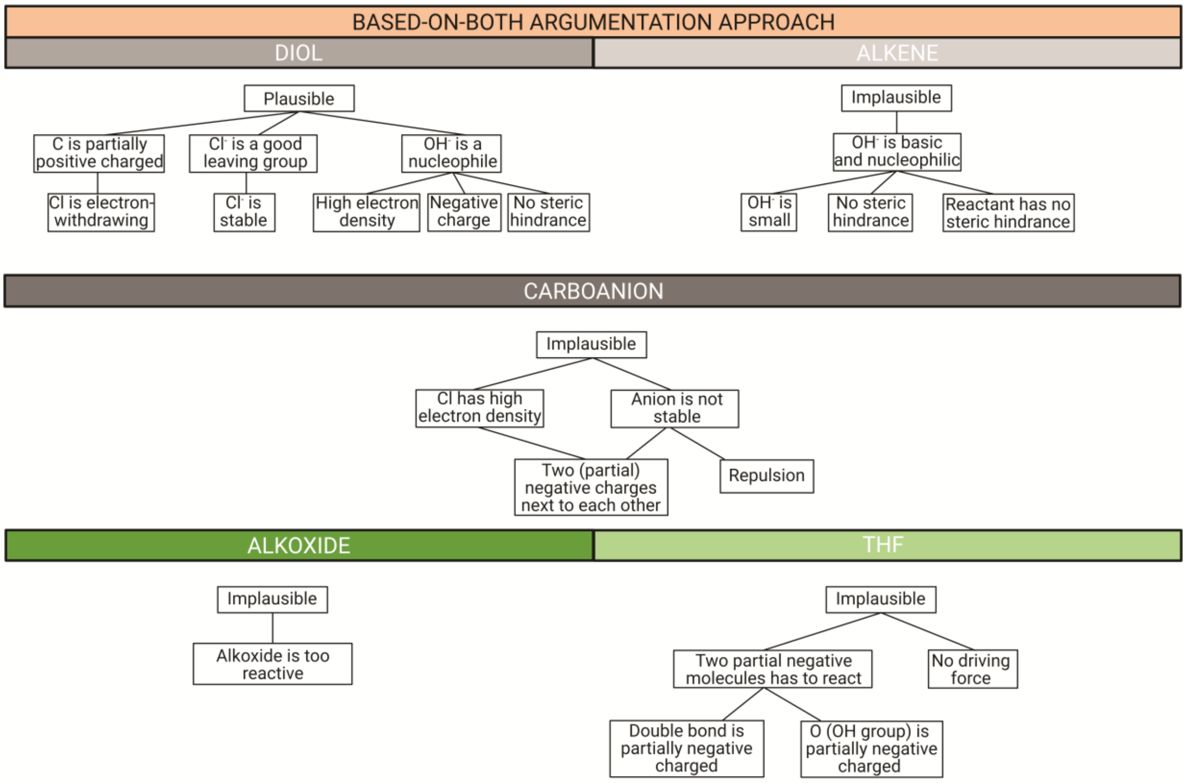


Fig. 30: Rhonda's argumentation pattern of the based-on-both argumentation approach.

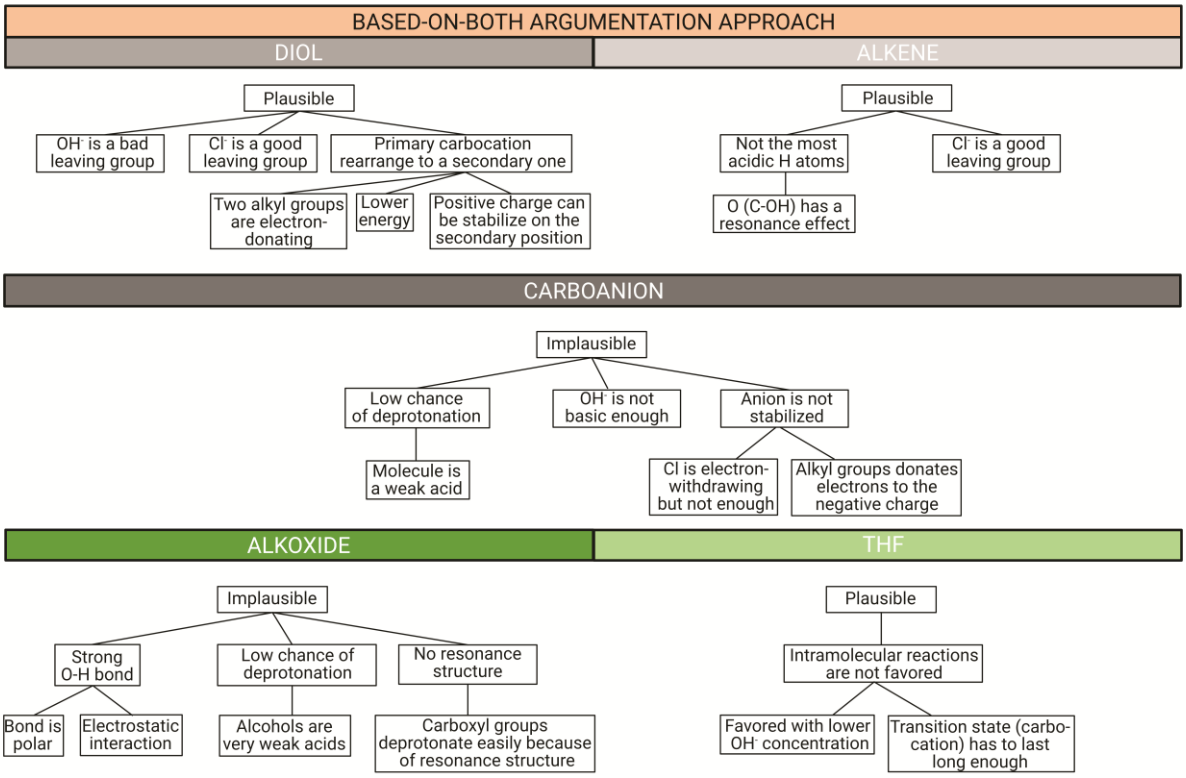


Fig. 31: Ronnie's argumentation pattern of the based-on-both argumentation approach.

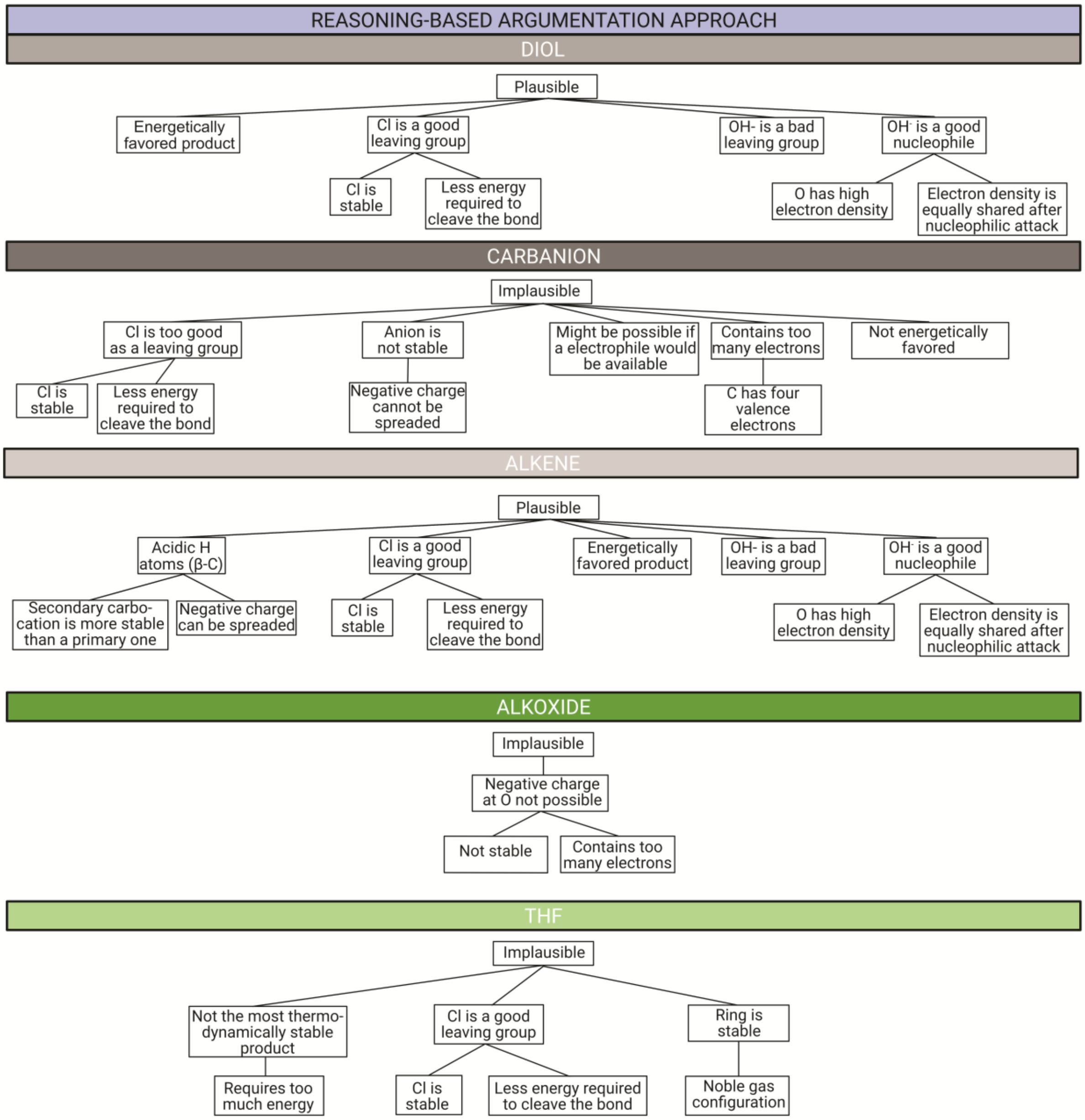


Fig. 32: Alex' argumentation pattern of the reasoning-based argumentation approach.

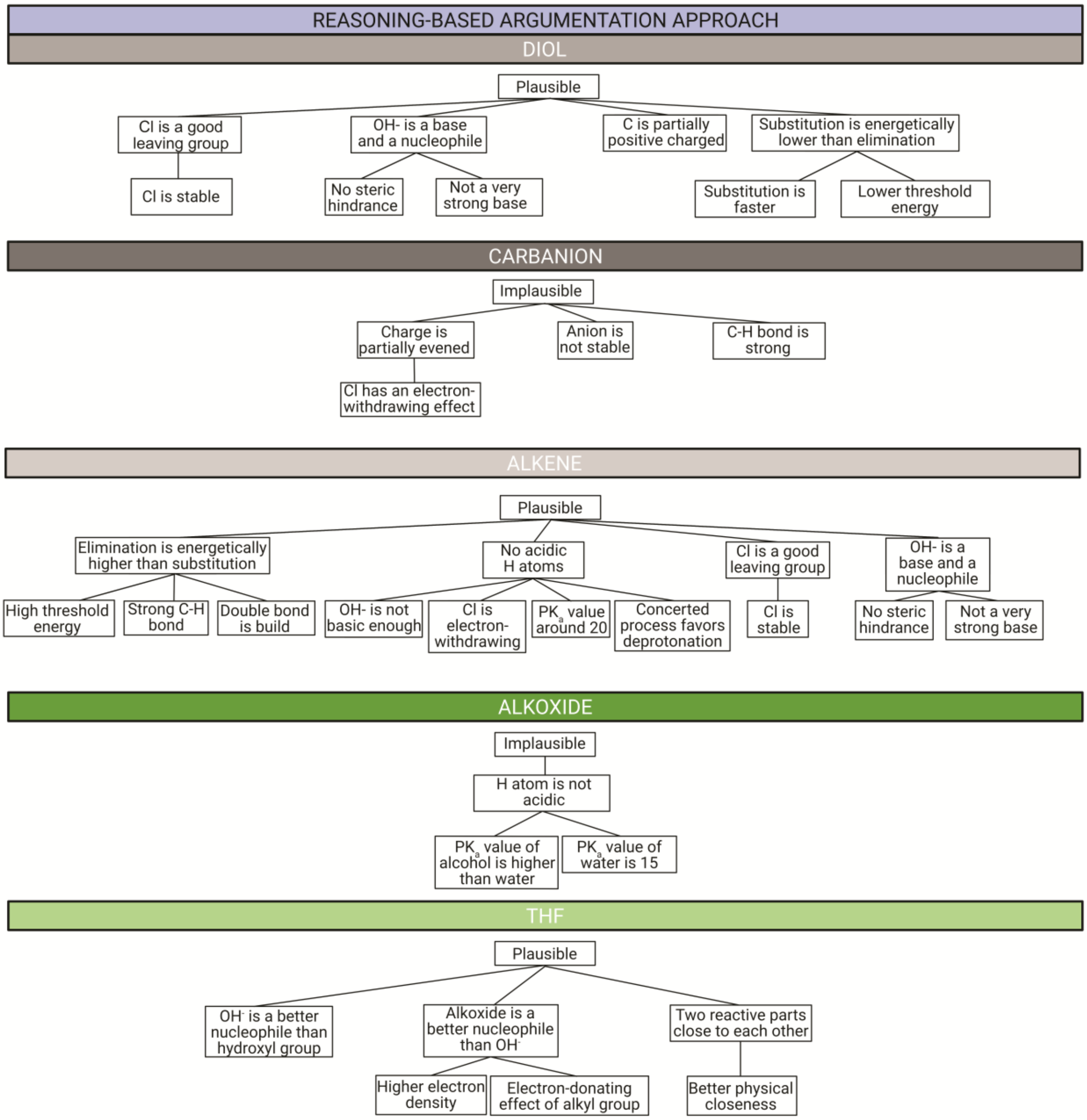


Fig. 33: Lily's argumentation pattern of the reasoning-based argumentation approach.

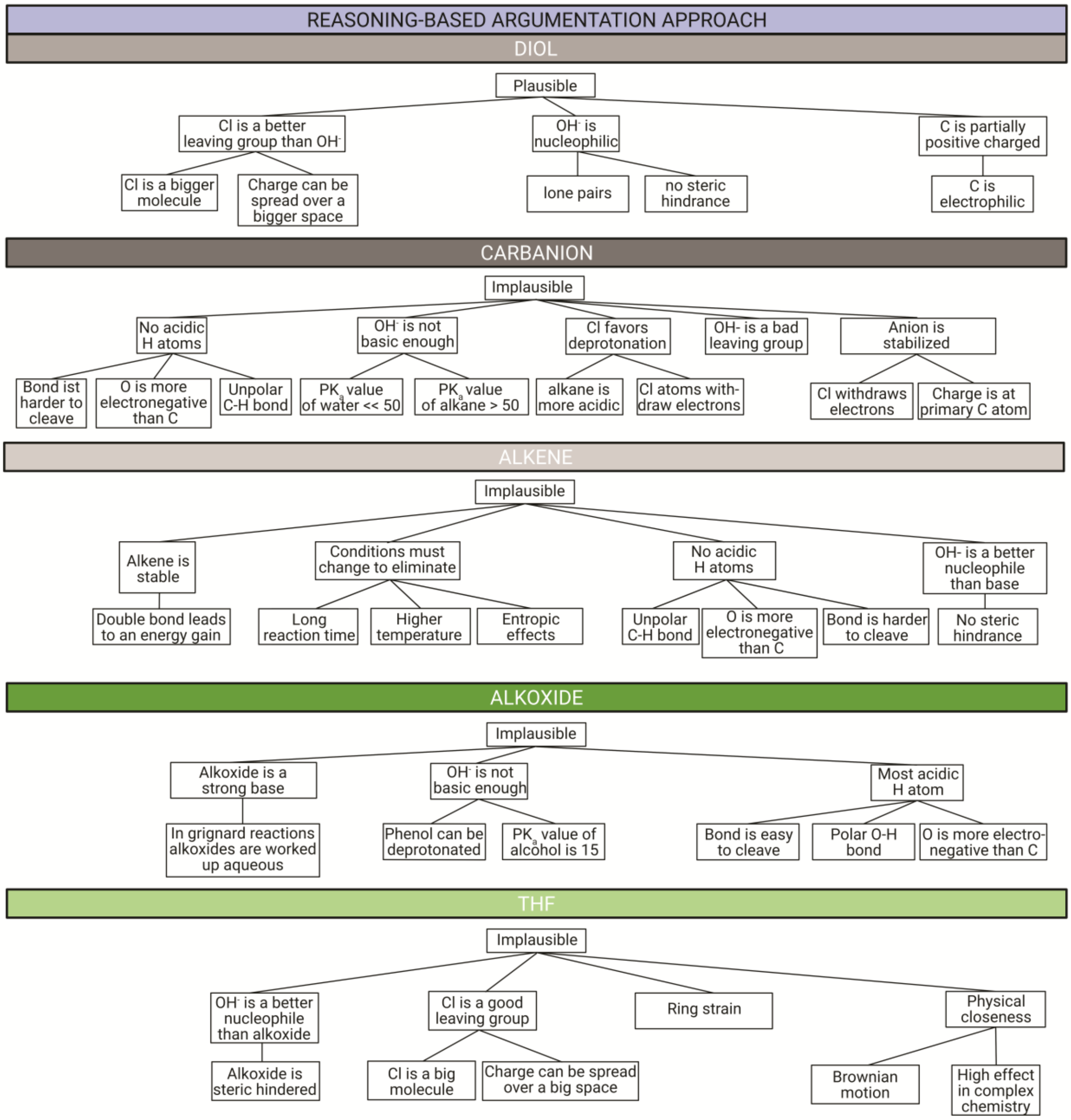


Fig. 34: Jay's argumentation pattern of the reasoning-based argumentation approach.

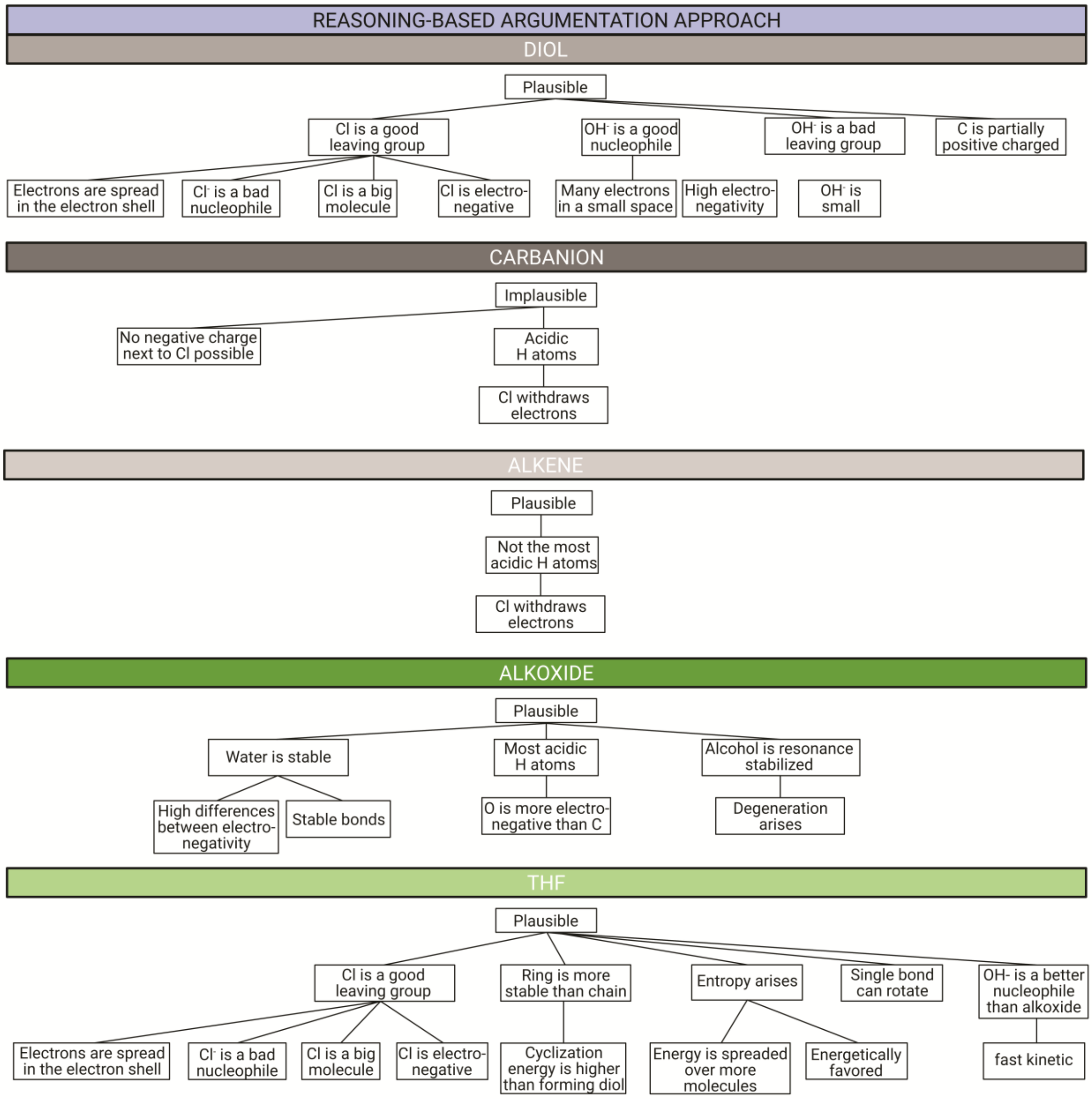


Fig. 35: Pepper's argumentation pattern of the reasoning-based argumentation approach.

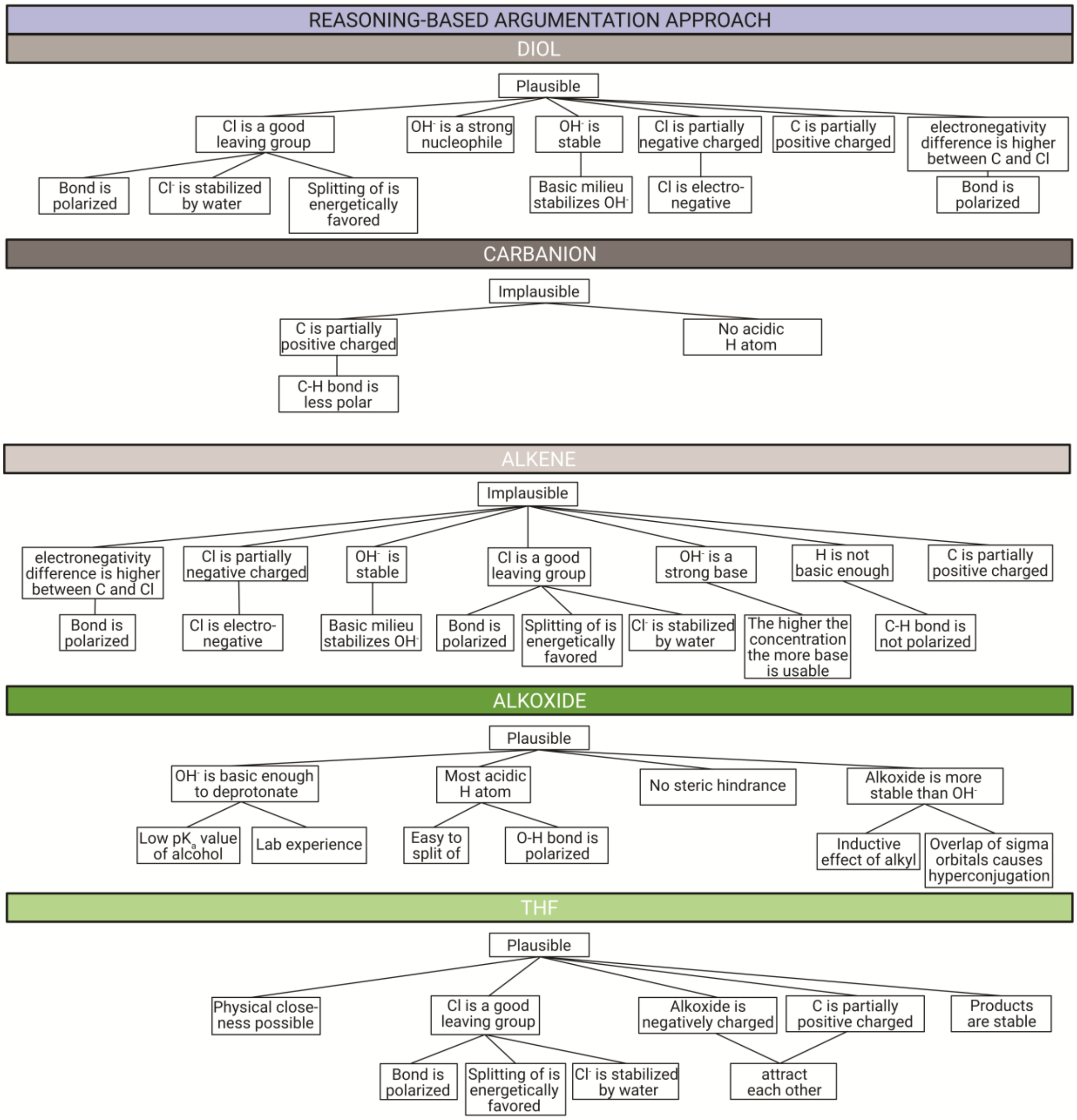


Fig. 36: Luke’s argumentation pattern of the reasoning-based argumentation approach.

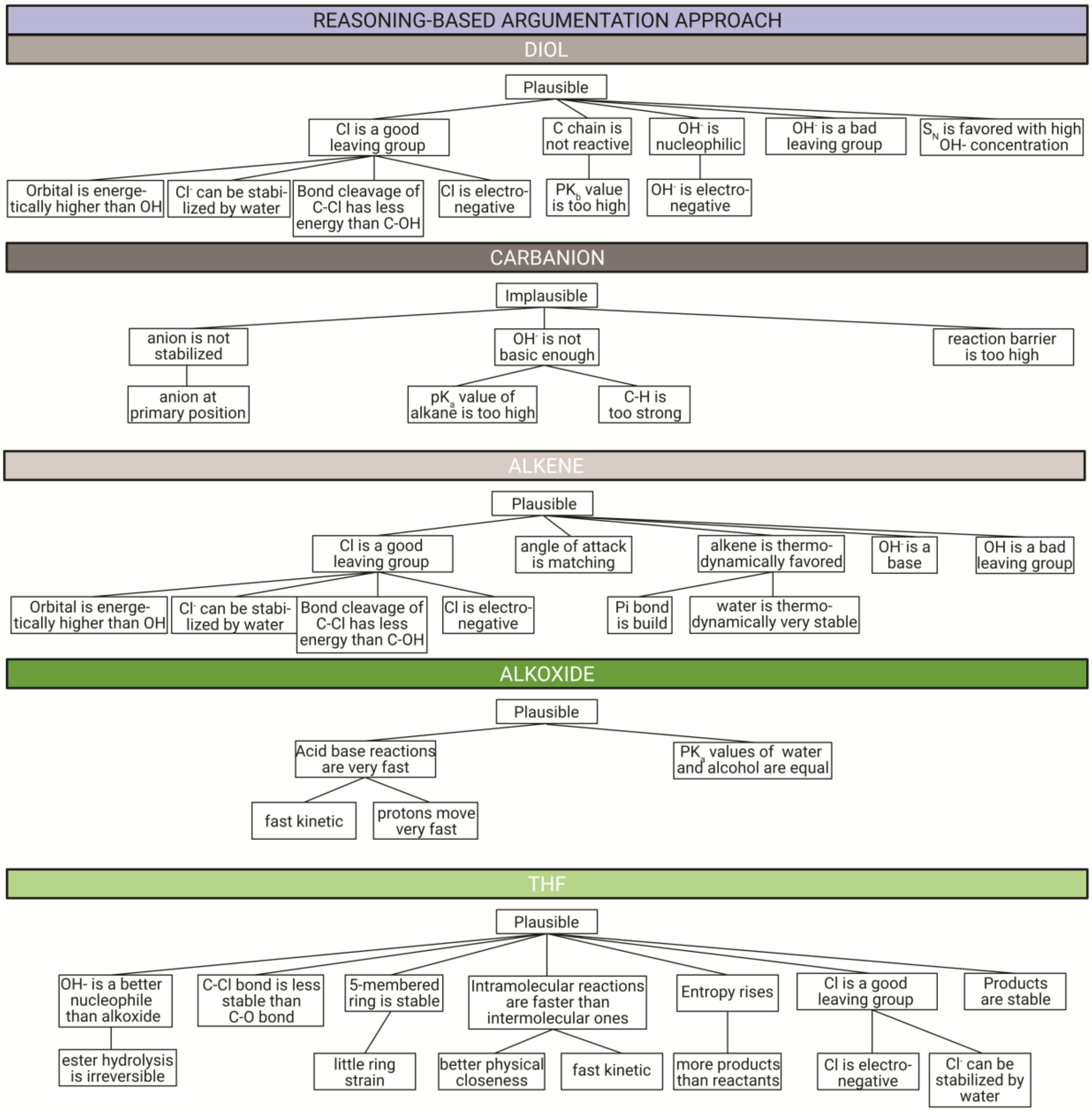


Fig. 37: Cameron's argumentation pattern of the reasoning-based argumentation approach.

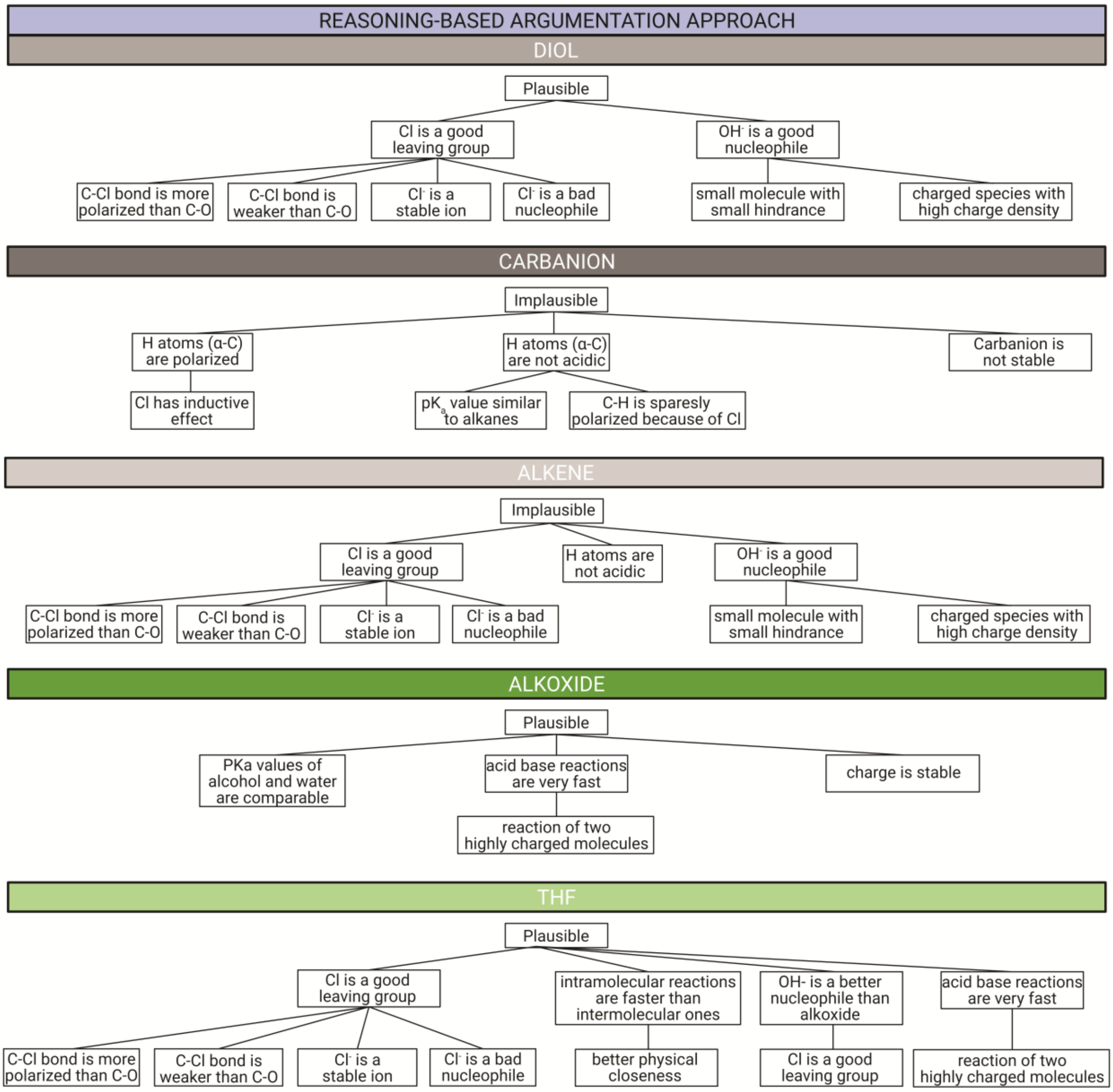


Fig. 38: Charlie's argumentation pattern of the reasoning-based argumentation approach.

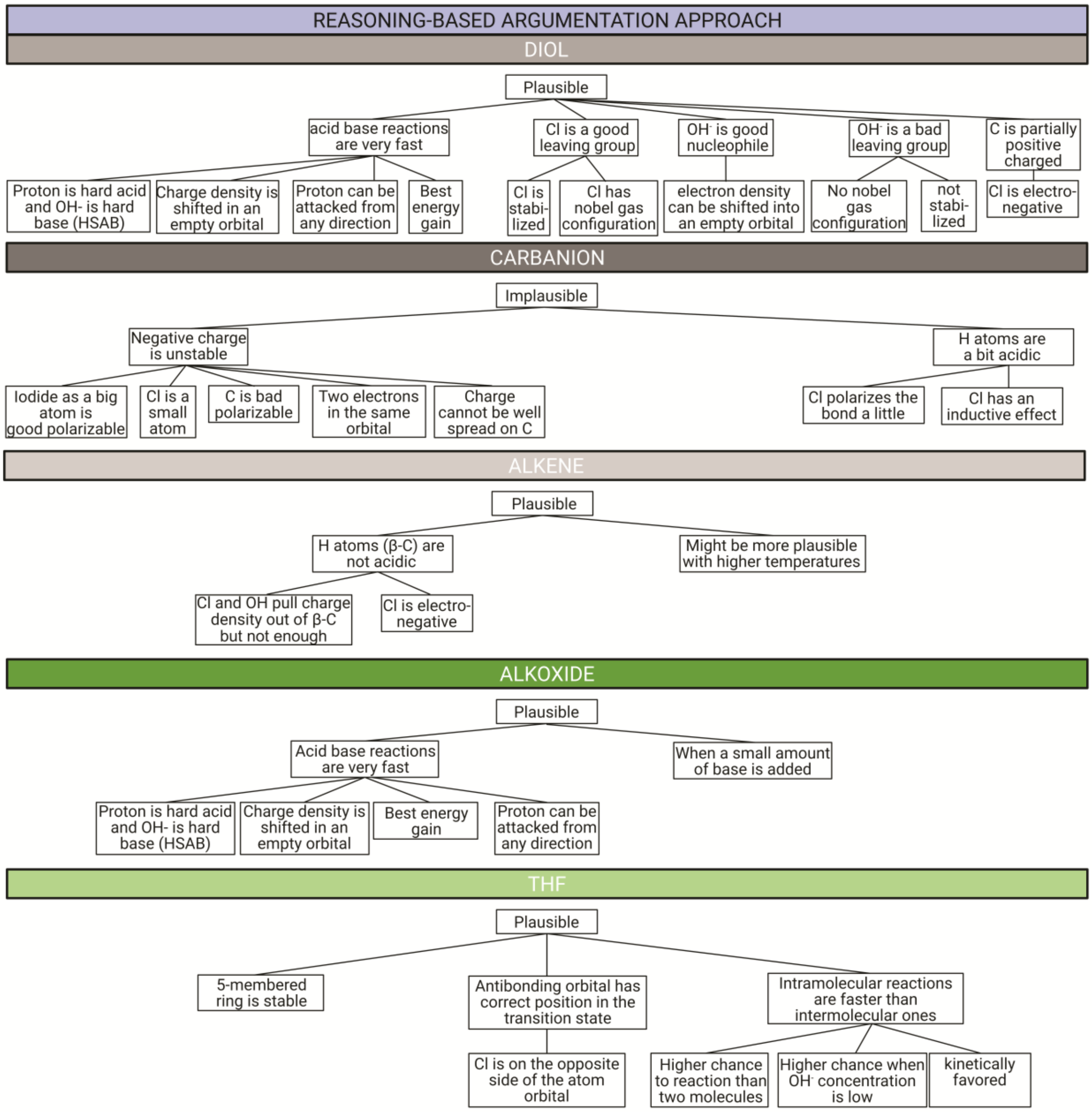


Fig. 39: Phil’s argumentation pattern of the reasoning-based argumentation approach.

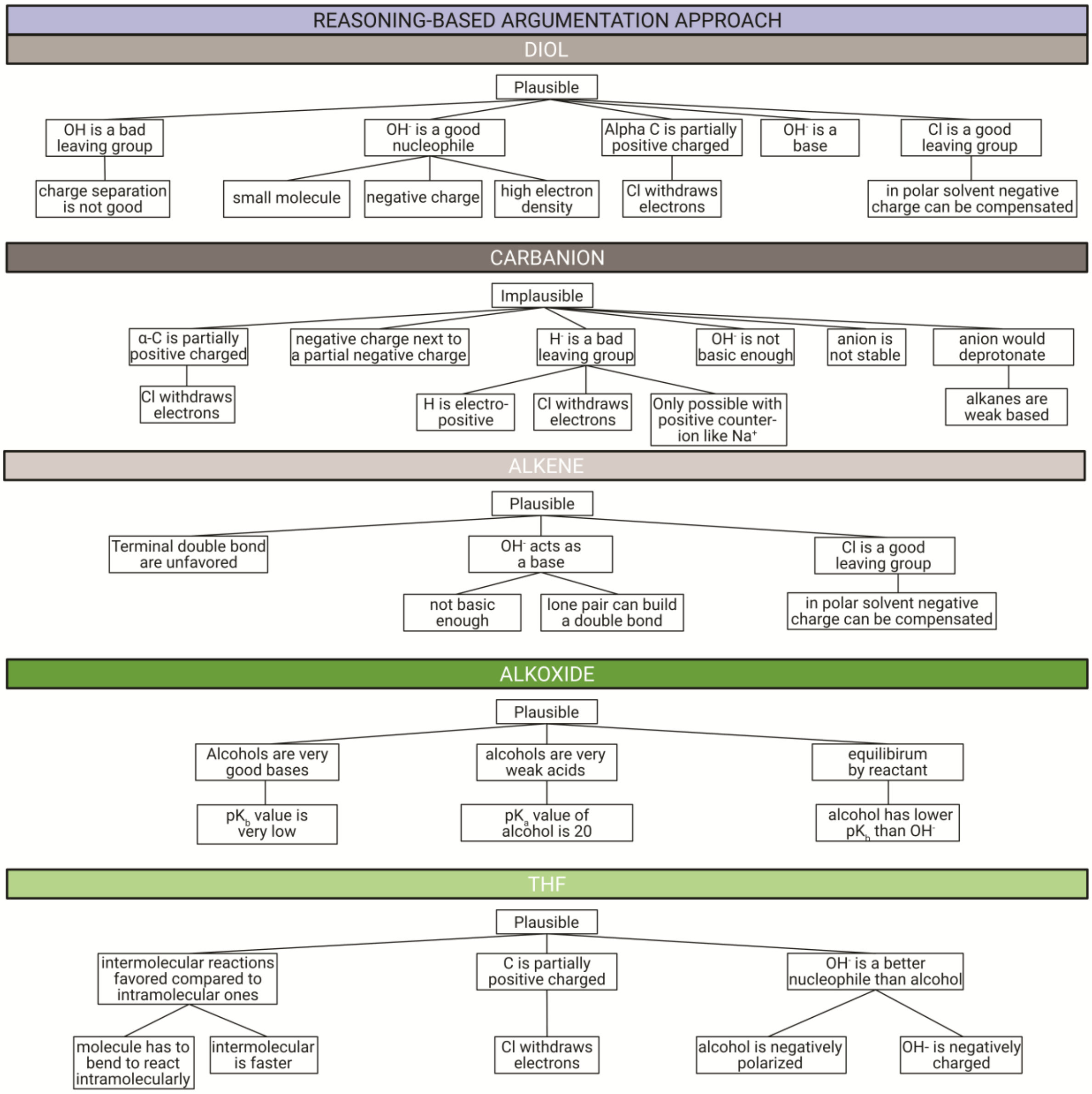


Fig. 40: Claire’s argumentation pattern of the reasoning-based argumentation approach.

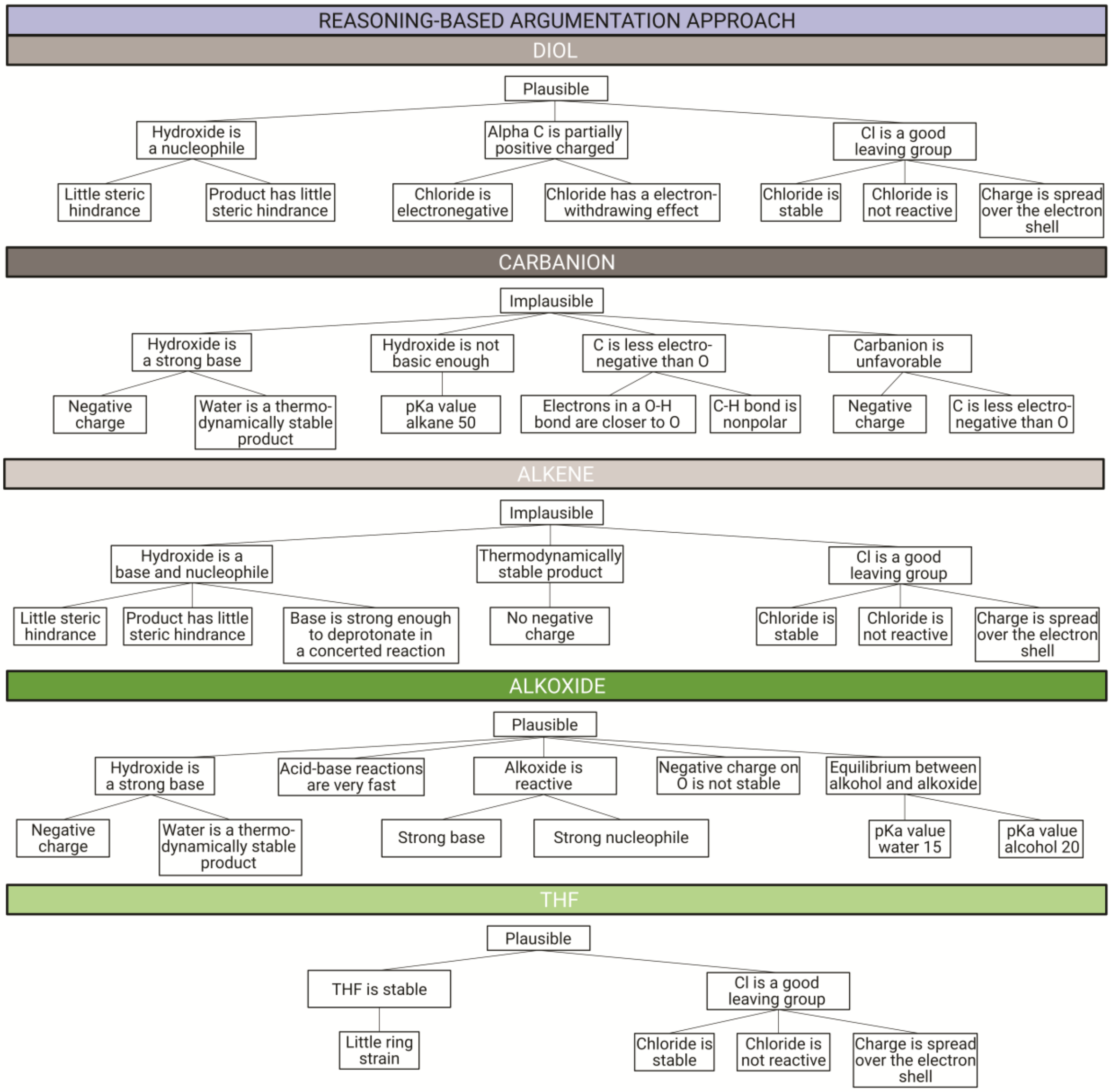


Fig. 41: Reuben's argumentation pattern of the reasoning-based argumentation approach.

6.3 Supporting Information Study II

6.3.1 Diagnostic Scaffold designed and used with the Software Qualtrics

Survey 1

Beginn des Blocks: Default Question Block

Q1 Welcome!
We (Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich) are grateful that you are part of the study.
If you have any questions, please feel free to contact Dr. Caspari (ira.caspari@tufts.edu).

Seitenumbruch

Q4 **User Code**
To maintain confidentiality in data collection, please create a code following the instructions below.

○ Third letter of the street of your address (*e.g.*, 20 In**G**ram Street, Queens) ________________

○ Last letter of the first name of a parent/guardian (if you have two parents/guardians, choose the older one) (*e.g.*, Mar**Y** or Be**N**) ________________

○ Second letter of the city/town/village your parent/guardian lives in (if you have two parents/guardians, choose the older one) (*e.g.*, M**E**dford) ________________

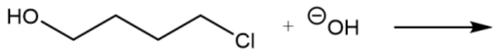
○ Date of birth: Day (two digits) (*e.g.*, 08) ________________

Seitenumbruch

Q5 Welcome!
In this survey you will be asked to build arguments about organic chemistry reactions. You will get specific support in how to build complete arguments and will practice this with several examples outside of organic chemistry to then apply it to organic chemistry. Please fill out this survey without any external support.

Seitenumbruch

Q6 **Task 1**
Which product is formed in the following reaction?
Draw the product *e.g.*, with Powerpoint, Word, with a drawing program like Chemdraw or take a picture of your drawing and upload the result as a file.



Q7 Build as many arguments as you can to justify your decision in as much detail as possible and in complete sentences.

○ Argument 1 ________________

○ Argument 2 ________________

○ Argument 3 ________________

○ Argument 4 ________________

○ Argument 5 ________________

Q8 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	very unconfident
How confident are you that you have drawn the correct reaction product?	○	○	○	○

Q9 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q10 **Exercise on Argumentation Patterns**

Building arguments and weighing these arguments are essential skills in science and everyday life. Often this process takes place unconsciously. For these reasons, we would like to explain and let you practice how to construct arguments in different contexts.

An argument should consist of three basic components:

The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Example: **If you want to appear taller, you should measure yourself in the morning**, **because people are on average three centimeters taller in the morning than in the evening**, **because their spinal discs lose fluid during the day and are compressed as a result.**

An argument, however, does not have to consist of just one mention of evidence or reasoning. An argument can be supported by <u>more than one</u> evidence and <u>more than one</u> line of reasoning.

Q11 Please read the following example about a students' judgement of the influence of the color red on bullfighting. After reading the example, you will be asked about what parts of this example can be categorizes as claim, evidence, and reasoning. Solutions can be found on the next page.

For bullfighting it is not important whether the torero is red, because the bull cannot recognize the color red, since it has receptors for blue and green only. Further the bull reacts rather to movement of the cloth than to color because it perceives the movement as a potential opponent due to the oppressive nature of the situation.

	Claim	Evidence	Reasoning	I don't know
For bullfighting it is not important whether the torero is red	○	○	○	○
because the bull cannot recognize the color red	○	○	○	○
since it has receptors for blue and green only.	○	○	○	○
Further the bull reacts rather to movement of the cloth than to color	○	○	○	○
because it perceives the movement as a potential opponent due to the oppressive nature of the situation.	○	○	○	○

Q12 Please read the following example about a students' judgement of an organic chemistry reaction step. After reading the example, you will be asked about what parts of this example can be categorizes as claim, evidence, and reasoning. Solutions can be found on the next page.

I think this reaction step is plausible because tosylate is a good leaving group and the product is very stable.

	Claim	Evidence	Reasoning	I don't know
I think this reaction step is plausible	○	○	○	○
because tosylate is a good leaving group	○	○	○	○
the product is very stable.	○	○	○	○

Q13 Please read the following example about a students' judgement of an organic chemistry reaction step. After reading the example, you will be asked about what parts of this example can be categorizes as claim, evidence, and reasoning. Solutions can be found on the next page.

The formation of a carbanion is implausible because the negative charge is unstable. This is because, that the carbon atom is a small molecule and poorly polarizable. It is also energetically unfavored that two electrons are in the same orbital.

	Claim	Evidence	Reasoning	I don't know
The formation of a carbanion is implausible	○	○	○	○
because the negative charge is unstable	○	○	○	○
This is because, that the carbon atom is a small molecule	○	○	○	○
and poorly polarizable	○	○	○	○
It is also energetically unfavored that two electrons are in the same orbital.	○	○	○	○

Seitenumbruch

Q14
Solutions to the previous exercise on argumentation patterns

Solution: For bullfighting it is not important whether the torero is red (Claim), because the bull cannot recognize the color red (Evidence), since it has receptors for blue and green only (Reasoning). and Further the bull reacts rather to the movement of the cloth than to color (Evidence), because it perceives the movement as a potential opponent due to the oppressive nature of the situation (Reasoning).

For bullfighting it is not important whether the torero is red	●	○	○	○
because the bull cannot recognize the color red	○	●	○	○
since it has receptors for blue and green only.	○	○	●	○
Further the bull reacts rather to movement of the cloth than to color	○	●	○	○
because it perceives the movement as a potential opponent due to the oppressive nature of the situation.	○	○	●	○

In the first exercise, a claim with two evidence statements, each supported by reasoning, was presented as a complete argument. This demonstrates that arguments can include several evidence and reasoning statements.

Solution: I think this reaction step is plausible (Claim) because tosylate is a good leaving group (Evidence) and the product is very stable (Evidence).

	Claim	Evidence	Reasoning	I don't know
I think this reaction step is plausible	●	○	○	○
because tosylate is a good leaving group	○	●	○	○
the product is very stable.	○	●	○	○

In the second exercise a claim was made that was supported by two evidence statements. However, reasoning for the evidence is missing. To complete the argument, it should have been added why tosylate is a good leaving group and why the product of the reaction is very stable. Here, one possible justification would be that tosylate is stabilized by resonance and the product (in an energy diagram) is very low in energy.

Solution: The formation of a carbanion is implausible (Claim) because the negative charge is unstable (Evidence). This is because, that the carbon atom is a small molecule (Reasoning) and poorly polarizable (Reasoning). It is also energetically unfavored that two electrons are in the same orbital (Reasoning).

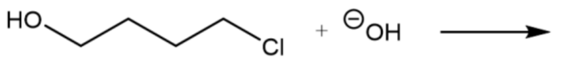
	Claim	Evidence	Reasoning	I don't know
The formation of a carbanion is implausible	●	○	○	○
because the negative charge is unstable	○	●	○	○
This is because, that the carbon atom is a small molecule	○	○	●	○
and poorly polarizable	○	○	●	○
It is also energetically unfavored that two electrons are in the same orbital.	○	○	●	○

In the third exercise a claim was made that was supported by evidence. This evidence was justified with three reasoning statements. Thus, a complete argument was constructed. Here it can be seen that arguments cannot only consist of several evidence but also of many reasoning statements.

Seitenumbruch

Q15 Exercise on Conceptual Understanding

Answer the prompts below and elaborate on your answers as detailed as possible. Please do not look up the answers on the internet or in the literature, but answers the prompts using your own knowledge only. Please do not leave any field blank. If you do not know the answer, write "I do not know" in the blank.



Answer the following tasks in as much detail as possible. In your answer do not only refer to the reactants of the reaction, but also to the products that you think form.
After typing your answers, take a photo or screenshot as you will not be able to see your answers anymore later in the survey, but you will still need them to answer further questions.

- ◯ Decide whether steric aspects need to be considered in the reaction and explain why you think so.

____

- ◯ Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule x has the highest and molecule y has the lowest pK_a)

____

- ◯ Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.

____

- ◯ Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.

____

- ◯ Determine whether you think there are any effects that stabilize your product compared to your reactants. If so, explain how the effect/s stabilize the product.

____

- ◯ Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.

____

- ◯ Decide whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.

____

- ◯ Decide whether the reaction is reactant- or product-favored from an energetic perspective (enthalpy). Explain your thinking.

____

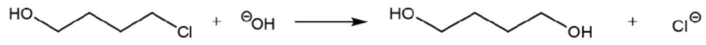
Q17 Did you take a picture or screenshot of your answers?

◯ Yes

◯ No

Seitenumbruch

Q16
Building arguments 1/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q18
Claim

◯ The reaction product is plausible

◯ The reaction product is implausible

Q19 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		

Argument 6		

Q22 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q23 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q24

Building arguments 2/4

Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q25

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q26 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q27 Answer the following question according to your personal experience.

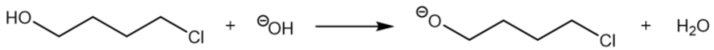
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q28 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q29
Building arguments 3/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q30

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q31 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		

Argument 4		
Argument 5		
Argument 6		

Q32 Answer the following question according to your personal experience.

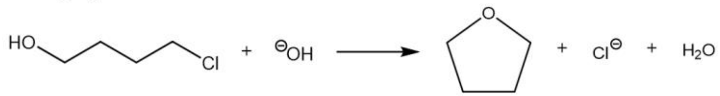
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q33 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q34
Building arguments 4/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q35

Claim

- ◯ The reaction product is plausible
- ◯ The reaction product is implausible

Q36 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q37 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	◯	◯	◯	◯

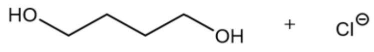
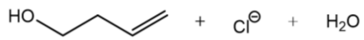
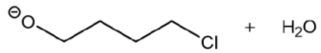
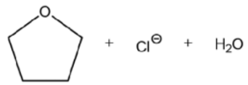
Q38 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	◯	◯	◯	◯

Seitenumbruch

Q39 **Decision-making**

Which product is formed in the following reaction? Select the reaction product/s. Only choose final product/s of the reaction, do not select intermediate/s.

- ☐ I stay with the reaction product I have formed before.
- ☐ 
- ☐ 
- ☐ 
- ☐ 

Seitenumbruch

Q40 **Evaluation**
Thank you very much for completing the chemistry portion of the survey. Before you submit the survey, we would like to ask you for your feedback about how useful you thought the argumentation training was. Please rate the following statements according to your personal experience.

	strongly agree	agree	neither agree nor disagree	Disagree	strongly disagree
The argumentation training helped me structure my arguments.	○	○	○	○	○
Prior to this survey, when answering tasks in organic chemistry, I rarely supported evidence with reasoning (*e.g.*, why leaving groups are good or bad)	○	○	○	○	○
The argumentation training helped me in answering the tasks.	○	○	○	○	○
I would use the argumentation pattern again when working on tasks in organic chemistry.	○	○	○	○	○
I would recommend the argumentation training to my fellow students.	○	○	○	○	○
I became more intensely involved with the tasks by building arguments.	○	○	○	○	○
I would have needed more help building arguments.	○	○	○	○	○
I was unsettled by the argumentation pattern.	○	○	○	○	○
Before the survey I did not know how arguments are constructed.	○	○	○	○	○
I would like to see building of arguments as part of the curriculum.	○	○	○	○	○
The questions about chemical concepts made me think of aspects that I had not included before.	○	○	○	○	○
The questions included key chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I found it difficult to apply these chemical concepts.	○	○	○	○	○
I was unsettled by the questions about chemical concepts.	○	○	○	○	○
I would have needed more help with questions about chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I became more intensely involved with the tasks by the questions on chemical concepts.	○	○	○	○	○

Seitenumbruch

Q41
Thank you for participating in the first part of the study! You will be notified in two weeks about the second part of the study.
Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich

Ende des Blocks: Default Question Block

6.3.2 Adapted Scaffold (ArgS group) designed and used with the Software Qualtrics

Survey 2 - Group 1

Beginn des Blocks: Default Question Block

Q3 Welcome!
We (Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich) are grateful that you are part of the study.
If you have any questions, please feel free to contact Dr. Caspari (ira.caspari@tufts.edu).

Seitenumbruch

Q5 **User Code**
To maintain confidentiality in data collection, please create a code following the instructions below.

○ Third letter of the street of your address (*e.g.*, 20 In**G**ram Street, Queens)
____

○ Last letter of the first name of a parent/guardian (if you have two parents/guardians, choose the older one) (*e.g.*, Mar**Y** or Be**N**)
____

○ Second letter of the city/town/village your parent/guardian lives in (if you have two parents/guardians, choose the older one) (*e.g.*, M**E**dford)
____

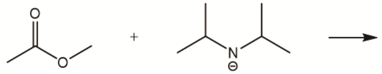
○ Date of birth: Day (two digits) (*e.g.*, 08)
____

Seitenumbruch

Q7 Welcome!
In this survey you will be asked to build arguments about organic chemistry reactions. You will get specific support in how to build complete arguments and will practice this with several examples outside of organic chemistry to then apply it to organic chemistry. Please fill out this survey without any external support.

Seitenumbruch

Q9 **Task 1**
Which product is formed in the following reaction?
Draw the product *e.g.*, with Powerpoint, Word, with a drawing program like Chemdraw or take a picture of your drawing and upload the result as a file.



Q11 Build as many arguments as you can to justify your decision in as much detail as possible and in complete sentences.

○ Argument 1 ________________________________
○ Argument 2 ________________________________
○ Argument 3 ________________________________
○ Argument 4 ________________________________
○ Argument 5 ________________________________

Q13 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	very unconfident
How confident are you that you have drawn the correct reaction product?	○	○	○	○

Q15 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q17 **Exercise on Argumentation Patterns**

Building arguments and weighing these arguments are essential skills in science and everyday life. Often this process takes place unconsciously. For these reasons, we would like to explain and let you practice how to construct arguments in different contexts.
An argument should consist of three basic components:
The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.
Example: **If you want to appear taller, you should measure yourself in the morning, because people are on average three centimeters taller in the morning than in the evening, because their spinal discs lose fluid during the day and are compressed as a result.**
An argument, however, does not have to consist of just one mention of evidence or reasoning. An argument can be supported by more than one evidence and more than one line of reasoning.

Q76
Please read the following example about a student's judgement of the use of canaries in mining. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Canaries serve as an "early warning system" in mining because they can give miners an early indication of whether carbon monoxide is in the air. Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning.

	Claim	Evidence	Reasoning	I don't know
Canaries serve as an "early warning system" in mining	○	○	○	○
because they can give miners an early indication of whether carbon monoxide is in the air.	○	○	○	○
Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Claim]

Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Evidence]

Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Reasoning]

Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [I don't know]

Q96 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: Canaries serve as an "early warning system" in mining (Claim), because they can give miners early indication of whether carbon monoxide is in the air (Evidence). Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning (Reasoning).

Q84
Please read the following example about a student's judgement of the protective function of birch barks. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The white color of the birch bark has a protective function, because it can regulate the temperature of the tree. The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously. This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark. In addition, the betulin in the white bark protects against bacteria and fungi, as it is antibacterial and water repellent.

	Claim	Evidence	Reasoning	I don't know
The white color of the birch bark has a protective function	○	○	○	○
because it can regulate the temperature of the tree	○	○	○	○
The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously.	○	○	○	○
This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark	○	○	○	○
In addition, the betulin in the white bark protects against bacteria and fungi	○	○	○	○
as it is antibacterial and water repellent.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Claim]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Evidence]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Reasoning]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [I don't know]

Q97 Here you can see the solution to compare it with your choice. Please do not change your initial answer.

Solution: The white color of the birch bark has a protective function (Claim), because it can regulate the temperature of the tree (Evidence). As the low winter sun causes the bark to be exposed to direct sunlight that increases the temperature of the bark tremendously (Reasoning). This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark (Reasoning). In addition, the betulin in the white bark protects against bacteria and fungi (Evidence), as it is antibacterial and water repellent (Reasoning).

Q85
Please read the following example about a student's judgement of bee stings. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Bees die after they sting as the resulting wound on their abdomen cannot heal, this is because the sting of the bee has 10 barbs that are torn out of the abdomen during the sting. In addition, after the sting, venom continues to be pumped through the muscles into the open wound.

	Claim	Evidence	Reasoning	I don't know
Bees die after they sting	○	○	○	○
as the resulting wound on their abdomen cannot heal	○	○	○	○
this is because the sting of the bee has 10 barbs that are torn out of the abdomen during the sting.	○	○	○	○
In addition, after the sting, venom continues to be pumped through the muscles into the open wound.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Claim]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Evidence]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Reasoning]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [I don't know]

Q98 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: Bees die after they sting (Claim) as the resulting wound of their abdomen cannot heal (Evidence), this is because the sting of a bee has 10 barbs that are torn out of the abdomen during the sting (Reasoning). In addition, after the sting, venom continues to be pumped through the muscles into the open wound (Reasoning).

Q86
Please read the following example about a student's judgement of the emergence of high and low tides. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The moon causes high and low tides because water moves differently depending on its position on earth's surface. This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull. Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.

	Claim	Evidence	Reasoning	I don't know
The moon causes high and low tides	○	○	○	○
because water moves differently depending on its position on earth's surface.	○	○	○	○
This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull.	○	○	○	○
Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Claim]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Evidence]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Reasoning]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [I don't know]

Q99 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The moon causes high and low tides (Claim) because water moves differently depending on its position on earth's surface (Evidence). This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull (Reasoning). Moreover, the actually small effect of acceleration leads to a high water pressure in the oceans by the collective migration of matter in large bodies of water (Reasoning).

Q94
Please read the following example about a student's judgement of an organic chemistry reaction step. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$, as the intermediate formed in the S_N1 reaction is stabilized, and because chloride is a good leaving group.

	Claim	Evidence	Reasoning	I don't know
In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$	○	○	○	○
as the intermediate formed in the S_N1 reaction is stabilized	○	○	○	○
and because chloride is a good leaving group.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Claim]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Evidence]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Reasoning]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [I don't know]

Q100 Here you can see the solution to compare it with your choice. Please do not change your initial answer.

Solution: **In a polar solvent, the S_N1 reaction of tert-butyl chloride with $AgNO_3$ proceeds faster than the S_N2 reaction of n-butyl chloride with $AgNO_3$ (Claim), as the intermediate formed in the SN1 reaction is stabilized (Evidence), and because chloride is a good leaving group (Evidence).**

Q95 Please read the following example about a student's judgement of an electrochemistry reaction. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The reaction of zinc ions with copper is reactant favored, as copper is not able to reduce zinc ions but zinc is able to reduce copper ions, since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent.

	Claim	Evidence	Reasoning	I don't know
The reaction of zinc ions with copper is reactant favored	○	○	○	○
as copper is not able to reduce zinc ions but zinc is able to reduce copper ions	○	○	○	○
since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Claim]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Evidence]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Reasoning]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [I don't know]

Q101 Here you can see the solution to compare it with your choice. Please do not change your initial answer.

Solution: **The reaction of zinc ions with copper is reactant favored (Claim), as copper is not able to reduce the zinc ions but zinc is able to reduce copper ions (Evidence), since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent (Reasoning).**

Q92 Please read the following example about a student's judgement of the polarity of the bond of hydrogen fluoride. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The bond in hydrogen fluoride is polar, due to the asymmetrical electron distribution in the bond. As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.

	Claim	Evidence	Reasoning	I don't know
The bond in hydrogen fluoride is polar	○	○	○	○
due to the asymmetrical electron distribution in the bond	○	○	○	○
As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Claim]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Evidence]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Reasoning]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [I don't know]

Q102 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The bond in hydrogen fluoride is polar (Claim), due to the asymmetrical electron distribution in the bond (Evidence). As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen (Reasoning).

Q93
Please read the following example about a student's judgement of ampholytes. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.
Water is an ampholyte since it reacts as an acid and as a base. Because oxygen has a high electronegativity and can stabilize charge well, water can act as a proton donor. Further, the lone pair of the oxygen can abstract a proton from another molecule.

	Claim	Evidence	Reasoning	I don't know
Water is an ampholyte	○	○	○	○
since it reacts as an acid and as a base	○	○	○	○
Because oxygen has a high electronegativity	○	○	○	○
and can stabilize charge well, water can act as a proton donor.	○	○	○	○
Further, the lone pair of the oxygen can abstract a proton from another molecule.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Claim]

Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Evidence]

Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Reasoning]

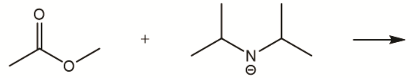
Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [I don't know]

Q103 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: Water is an ampholyte (Claim), since it reacts as an acid and as a base (Evidence). Because oxygen has a high electronegativity (Reasoning) and can stabilize negative charge well, water can act as a proton donor (Reasoning). Further, the lone pair on the oxygen can abstract a proton from another molecule (Reasoning).

Seitenumbruch

Q19 **Exercise on Conceptual Understanding**

Answer the prompts below and <u>elaborate on your answers as detailed as possible</u>. Please do not look up the answers on the internet or in the literature, but answers the prompts using your own knowledge only. Please do not leave any field blank. If you do not know the answer, write "I do not know" in the blank.



Answer the following tasks in as much detail as possible. In your answer do not only refer to the reactants of the reaction, but also to the products that you think form.
<u>After typing your answers, take a photo or screenshot as you will not be able to see your answers anymore later in the survey, but you will still need them to answer further questions.</u>

○ Decide whether steric aspects need to be considered in the reaction and explain why you think so.

____

○ Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule x has the highest and molecule y has the lowest pK_a)

____

○ Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.

____

○ Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.

____

○ Determine whether you think there are any effects that stabilize your product compared to your reactants. If so, explain how the effect/s stabilize the product.

____

○ Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.

____

○ Decide whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.

____

○ Decide whether the reaction is reactant- or product-favored from an energetic perspective (enthalpy). Explain your thinking.

____

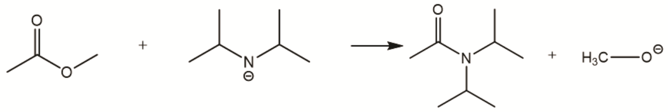
Q21 Did you take a picture or screenshot of your answers?

○ Yes

○ No

Seitenumbruch

Q23
Building arguments 1/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q25
The **claim** represents the position for which one argues. The **evidence** supports the claim and consists of data on which the claim is based. The **reasoning** justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q27 Build as many arguments as you can.

	Evidence (Support of the claim)	**Reasoning** (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		

Argument 4		
Argument 5		
Argument 6		

Q29 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

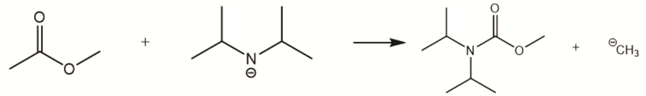
Q31 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q33

Building arguments 2/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q35

The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q37 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		

Argument 5		
Argument 6		

Q39 Answer the following question according to your personal experience.

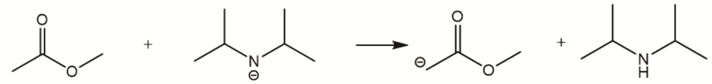
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q41 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q43
Building arguments 3/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q45
The **claim** represents the position for which one argues. The **evidence** supports the claim and consists of data on which the claim is based. The **reasoning** justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q47 Build as many arguments as you can.

	Evidence (Support of the claim)	**Reasoning** (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		

Argument 5		
Argument 6		

Q49 Answer the following question according to your personal experience.

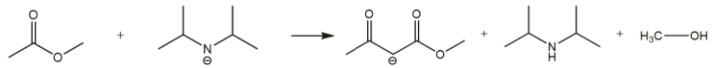
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q51 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q53
Building arguments 4/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q55
The **claim** represents the position for which one argues. The **evidence** supports the claim and consists of data on which the claim is based. The **reasoning** justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q57 Build as many arguments as you can.

	Evidence (Support of the claim)	**Reasoning** (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		

Argument 5		
Argument 6		

Q59 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

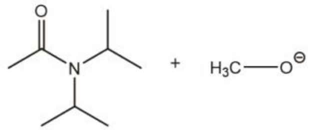
Q61 Answer the following question according to your personal experience.

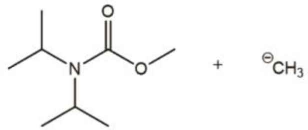
	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

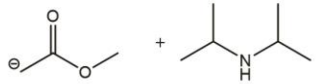
Seitenumbruch

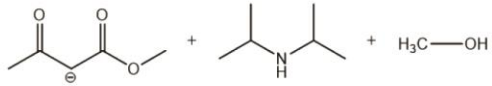
Q63 **Decision-making**

Which product is formed in the following reaction? Select the reaction product/s. Only choose final product/s of the reaction, do not select intermediate/s.

- [ ] I stay with the reaction product I have formed before.
- [ ] 

- [ ] 

- [ ] 

- [ ] 

Seitenumbruch

Q65 Evaluation
Thank you very much for completing the chemistry portion of the survey. Before you submit the survey, we would like to ask you for your feedback about how useful you thought the argumentation training was.
Please rate the following statements according to your personal experience.

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
The argumentation training helped me structure my arguments.	◯	◯	◯	◯	◯
Prior to this survey, when answering tasks in organic chemistry, I rarely supported evidence with reasoning (*e.g.*, why leaving groups are good or bad)	◯	◯	◯	◯	◯
The argumentation training helped me in answering the tasks.	◯	◯	◯	◯	◯
I would use the argumentation pattern again when working on tasks in organic chemistry.	◯	◯	◯	◯	◯
I would recommend the argumentation training to my fellow students.	◯	◯	◯	◯	◯
I became more intensely involved with the tasks by building arguments.	◯	◯	◯	◯	◯
I would have needed more help building arguments.	◯	◯	◯	◯	◯
I was unsettled by the argumentation pattern.	◯	◯	◯	◯	◯
Before the survey I did not know how arguments are constructed.	◯	◯	◯	◯	◯
I would like to see building of arguments as part of the curriculum.	◯	◯	◯	◯	◯
The questions about chemical concepts made me think of aspects that I had not included before.	◯	◯	◯	◯	◯

The questions included key chemical concepts (*e.g.*, nucleophilicity).	◯	◯	◯	◯	◯
I found it difficult to apply these chemical concepts.	◯	◯	◯	◯	◯
I was unsettled by the questions about chemical concepts.	◯	◯	◯	◯	◯
I would have needed more help with questions about chemical concepts (*e.g.*, nucleophilicity).	◯	◯	◯	◯	◯
I became more intensely involved with the tasks by the questions on chemical concepts.	◯	◯	◯	◯	◯

Q68 Here you have the possibility to give feedback.

____
____
____
____
____

Seitenumbruch

Q67
Thank you for participating in the second part of the study!
Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich

Ende des Blocks: Default Question Block

6.3.3 Adapted Scaffold (ConS group) designed and used with the Software Qualtrics

Survey 2 - Group 2

Beginn des Blocks: Default Question Block

Q3 Welcome!
We (Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich) are grateful that you are part of the study.
If you have any questions, please feel free to contact Dr. Caspari (ira.caspari@tufts.edu).

Seitenumbruch

Q5 **User Code**
To maintain confidentiality in data collection, please create a code following the instructions below.

○ Third letter of the street of your address (*e.g.*, 20 In**G**ram Street, Queens)
____

○ Last letter of the first name of a parent/guardian (if you have two parents/guardians, choose the older one) (*e.g.*, Mar**Y** or Be**N**)
____

○ Second letter of the city/town/village your parent/guardian lives in (if you have two parents/guardians, choose the older one) (*e.g.*, M**E**dford)
____

○ Date of birth: Day (two digits) (*e.g.*, 08)
____

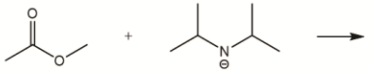
Seitenumbruch

Q7 Welcome!
In this survey you will be asked to build arguments about organic chemistry reactions. You will get specific support in how to build complete arguments and will practice this with several examples outside of organic chemistry to then apply it to organic chemistry. Please fill out this survey without any external support.

Seitenumbruch

Page 1 of 22

Q9 **Task 1**
Which product is formed in the following reaction?
Draw the product *e.g.*, with Powerpoint, Word, with a drawing program like Chemdraw or take a picture of your drawing and upload the result as a file.



Q11 Build as many arguments as you can to justify your decision in as much detail as possible and in complete sentences.

○ Argument 1 ________________________________

○ Argument 2 ________________________________

○ Argument 3 ________________________________

○ Argument 4 ________________________________

○ Argument 5 ________________________________

Q13 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	very unconfident
How confident are you that you have drawn the correct reaction product?	○	○	○	○

Q15 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Page 2 of 22

Q17 **Exercise on Argumentation Patterns**

Building arguments and weighing these arguments are essential skills in science and everyday life. Often this process takes place unconsciously. For these reasons, we would like to explain and let you practice how to construct arguments in different contexts.
An argument should consist of three basic components:
The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.
Example: **If you want to appear taller, you should measure yourself in the morning, because people are on average three centimeters taller in the morning than in the evening, because their spinal discs lose fluid during the day and are compressed as a result.**
An argument, however, does not have to consist of just one mention of evidence or reasoning. An argument can be supported by more than one evidence and more than one line of reasoning.

Q19
Please read the following example about a student's judgement of the emergence of high and low tides. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The moon causes high and low tides because water moves differently depending on its position on earth's surface. This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull. Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.

	Claim	Evidence	Reasoning	I don't know
The moon causes high and low tides	○	○	○	○
because water moves differently depending on its position on earth's surface.	○	○	○	○
This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull.	○	○	○	○
Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Claim]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Evidence]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Reasoning]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [I don't know]

Q21 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The moon causes high and low tides (Claim) because water moves differently depending on its position on earth's surface (Evidence). This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull (Reasoning). Moreover, the actually small effect of acceleration leads to a high water pressure in the oceans by the collective migration of matter in large bodies of water (Reasoning).

Q23
Please read the following example about a student's judgement of an organic chemistry reaction step. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$, as the intermediate formed in the S_N1 reaction is stabilized, and because chloride is a good leaving group.

	Claim	Evidence	Reasoning	I don't know
In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$	○	○	○	○
as the intermediate formed in the S_N1 reaction is stabilized	○	○	○	○
and because chloride is a good leaving group.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Claim]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Evidence]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Reasoning]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [I don't know]

Q25 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: **In a polar solvent, the S_N1 reaction of tert-butyl chloride with $AgNO_3$ proceeds faster than the S_N2 reaction of n-butyl chloride with $AgNO_3$ (Claim), as the intermediate formed in the S_N1 reaction is stabilized (Evidence), and because chloride is a good leaving group (Evidence).**

Q27 Please read the following example about a student's judgement of the polarity of the bond of hydrogen fluoride. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The bond in hydrogen fluoride is polar, due to the asymmetrical electron distribution in the bond. As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.

	Claim	Evidence	Reasoning	I don't know
The bond in hydrogen fluoride is polar	○	○	○	○
due to the asymmetrical electron distribution in the bond	○	○	○	○
As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Claim]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Evidence]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Reasoning]

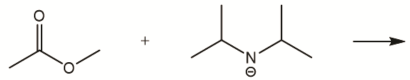
Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [I don't know]

Q29 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: **The bond in hydrogen fluoride is polar (Claim), due to the asymmetrical electron distribution in the bond (Evidence). As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen (Reasoning).**

Seitenumbruch

Q31 **Exercise on Conceptual Understanding**

Answer the prompts below and elaborate on your answers as detailed as possible. Please do not look up the answers on the internet or in the literature, but answers the prompts using your own knowledge only. Please do not leave any field blank. If you do not know the answer, write "I do not know" in the blank.



Answer the following tasks in as much detail as possible. In your answer do not only refer to the reactants of the reaction, but also to the products that you think form.
After typing your answers, take a photo or screenshot as you will not be able to see your answers anymore later in the survey, but you will still need them to answer further questions.

○ Decide whether steric aspects need to be considered in the reaction and explain why you think so.

○ Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule x has the highest and molecule y has the lowest pK_a)

○ Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.

○ Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.

○ Determine whether you think there are any effects that stabilize your product compared to your reactants. If so, explain how the effect/s stabilize the product.

○ Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.

○ Decide whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.

○ Decide whether the reaction is reactant- or product-favored from an energetic perspective (enthalpy). Explain your thinking.

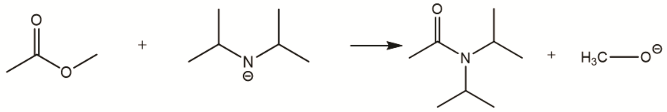
Q33 Did you take a picture or screenshot of your answers?

○ Yes

○ No

Seitenumbruch

Q35
Building arguments 1/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q82
In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.
Nucleophilicity/Electrophilicity
Nucleophilic centers: lone pairs at nitrogen and oxygen
Most electrophilic center: partial positively charged carbonyl C
Acidity/Basicity
Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.
Base: In terms of Lewis capable of donating an electron pair. Most basic atom is the nitrogen in the reactant.
Acid-base-reactions are usually faster than other reaction types.
Steric Aspects
The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.
Entropic Aspects
Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.
Electronic Aspects
Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.
Stability
A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.
pK_a Values
CH3OH approx. 15, ester approx. 25, diisopropylamine approx. 40
Energetic Aspects
Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q37

Claim

- ◯ The reaction product is plausible
- ◯ The reaction product is implausible

Q39 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q41 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	◯	◯	◯	◯

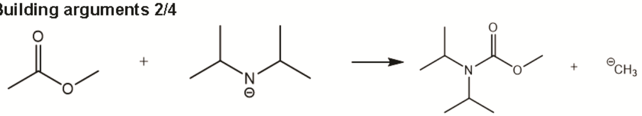
Q43 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	◯	◯	◯	◯

Seitenumbruch

Q45

Building arguments 2/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q83

In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.

Nucleophilicity/Electrophilicity

Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon
Most electrophilic center: partial positively charged carbonyl C

Acidity/Basicity

Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.
Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant and the negatively charged carbon in the product.
Acid-base-reactions are usually faster than other reaction types.

Steric Aspects

The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.

Entropic Aspects

Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.

Electronic Aspects

Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.

Stability

A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values

CH4 approx. 56, ester approx. 25, diisopropylamine approx. 40

Energetic Aspects

Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q47

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q49 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q51 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q53 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q55

Building arguments 3/4

Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q84

In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.

Nucleophilicity/Electrophilicity

Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon

Most electrophilic center: partial positively charged carbonyl C

Acidity/Basicity

Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.

Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant.

Acid-base-reactions are usually faster than other reaction types.

Steric Aspects

The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.

Entropic Aspects

Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.

Electronic Aspects

Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2

Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.

Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.

Stability

A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values

ester approx. 25, diisopropylamine approx. 40

Energetic Aspects

Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q57

Claim

- ◯ The reaction product is plausible
- ◯ The reaction product is implausible

Q59 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q61 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	◯	◯	◯	◯

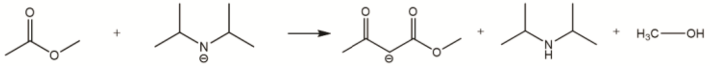
Q63 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	◯	◯	◯	◯

Seitenumbruch

Q65

Building arguments 4/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q85

In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.

Nucleophilicity/Electrophilicity
Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon
Most electrophilic center: partial positively charged carbonyl C

Acidity/Basicity
Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group and in the further reaction the proton at the carbon atom between the two carbonyl groups.
Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant and the negatively charged oxygen of methoxide as an intermediate.
Acid-base-reactions are usually faster than other reaction types.

Steric Aspects
The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.

Entropic Aspects
Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.

Electronic Aspects
Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates and the product show resonance.

Stability
A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values
ester approx. 25, diisopropylamine approx. 40, methyl acetoacetate approx. 12

Energetic Aspects
Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q67

Claim

- ◯ The reaction product is plausible
- ◯ The reaction product is implausible

Q69 Build as many arguments as you can.

	Evidence (Support of the claim)	**Reasoning** (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q71 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q73 Answer the following question according to your personal experience.

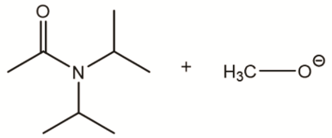
	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

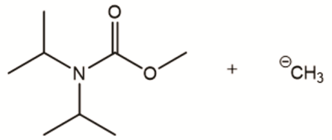
Seitenumbruch

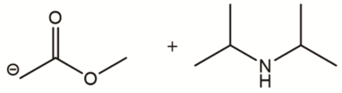
Q47 **Decision-making**

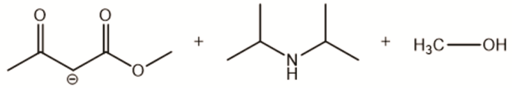
Which product is formed in the following reaction? Select the reaction product/s. Only choose final product/s of the reaction, do not select intermediate/s.

☐ I stay with the reaction product I have formed before.

☐ 

☐ 

☐ 

☐ 

Seitenumbruch

Q77 **Evaluation**
Thank you very much for completing the chemistry portion of the survey. Before you submit the survey, we would like to ask you for your feedback about how useful you thought the argumentation training was.
Please rate the following statements according to your personal experience.

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
The argumentation training helped me structure my arguments.	○	○	○	○	○
Prior to this survey, when answering tasks in organic chemistry, I rarely supported evidence with reasoning (*e.g.*, why leaving groups are good or bad)	○	○	○	○	○
The argumentation training helped me in answering the tasks.	○	○	○	○	○
I would use the argumentation pattern again when working on tasks in organic chemistry.	○	○	○	○	○
I would recommend the argumentation training to my fellow students.	○	○	○	○	○
I became more intensely involved with the tasks by building arguments.	○	○	○	○	○
I would have needed more help building arguments.	○	○	○	○	○
I was unsettled by the argumentation pattern.	○	○	○	○	○
Before the survey I did not know how arguments are constructed.	○	○	○	○	○
I would like to see building of arguments as part of the curriculum.	○	○	○	○	○
The questions about chemical concepts made me think of aspects that I had not included before.	○	○	○	○	○
The questions included key chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○

I found it difficult to apply these chemical concepts.	○	○	○	○	○
I was unsettled by the questions about chemical concepts.	○	○	○	○	○
I would have needed more help with questions about chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I became more intensely involved with the tasks by the questions on chemical concepts.	○	○	○	○	○

Q79 Here you have the possibility to give feedback.

____
____
____
____
____

Seitenumbruch

Q81
Thank you for participating in the second part of the study!
Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich

Ende des Blocks: Default Question Block

Survey 2 - Group 3

Beginn des Blocks: Default Question Block

Q3 Welcome!
We (Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich) are grateful that you are part of the study.
If you have any questions, please feel free to contact Dr. Caspari (ira.caspari@tufts.edu).

Seitenumbruch

Q5 **User Code**
To maintain confidentiality in data collection, please create a code following the instructions below.

○ Third letter of the street of your address (*e.g.*, 20 In**G**ram Street, Queens)

____

○ Last letter of the first name of a parent/guardian (if you have two parents/guardians, choose the older one) (*e.g.*, Mar**Y** or Be**N**)

____

○ Second letter of the city/town/village your parent/guardian lives in (if you have two parents/guardians, choose the older one) (*e.g.*, M**E**dford)

____

○ Date of birth: Day (two digits) (*e.g.*, 08)

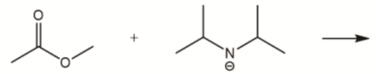
____

Seitenumbruch

Q7 Welcome!
In this survey you will be asked to build arguments about organic chemistry reactions. You will get specific support in how to build complete arguments and will practice this with several examples outside of organic chemistry to then apply it to organic chemistry. Please fill out this survey without any external support.

Seitenumbruch

Q9 **Task 1**
Which product is formed in the following reaction?
Draw the product *e.g.*, with Powerpoint, Word, with a drawing program like Chemdraw or take a picture of your drawing and upload the result as a file.



Q11 Build as many arguments as you can to justify your decision in as much detail as possible and in complete sentences.

○ Argument 1 ________________________________

○ Argument 2 ________________________________

○ Argument 3 ________________________________

○ Argument 4 ________________________________

○ Argument 5 ________________________________

Q13 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	very unconfident
How confident are you that you have drawn the correct reaction product?	○	○	○	○

Q15 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q17 **Exercise on Argumentation Patterns**

Building arguments and weighing these arguments are essential skills in science and everyday life. Often this process takes place unconsciously. For these reasons, we would like to explain and let you practice how to construct arguments in different contexts.
An argument should consist of three basic components:
The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.
Example: **If you want to appear taller, you should measure yourself in the morning, because people are on average three centimeters taller in the morning than in the evening, because their spinal discs lose fluid during the day and are compressed as a result.**
An argument, however, does not have to consist of just one mention of evidence or reasoning. An argument can be supported by more than one evidence and more than one line of reasoning.

Q19
Please read the following example about a student's judgement of the use of canaries in mining. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Canaries serve as an "early warning system" in mining because they can give miners an early indication of whether carbon monoxide is in the air. Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning.

	Claim	Evidence	Reasoning	I don't know
Canaries serve as an "early warning system" in mining	○	○	○	○
because they can give miners an early indication of whether carbon monoxide is in the air.	○	○	○	○
Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning.	○	○	○	○

Diese Frage anzeigen:
If Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Claim]
Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Evidence]
Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [Reasoning]
Or Please read the following example about a student's judgement of the use of canaries in mining. A... = Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning. [I don't know]

Q21 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: Canaries serve as an "early warning system" in mining (Claim), because they can give miners early indication of whether carbon monoxide is in the air (Evidence). Due to the small lung volume of the birds, even a small amount of carbon monoxide is sufficient for them to show typical symptoms of poisoning (Reasoning).

Q23
Please read the following example about a student's judgement of the protective function of birch barks. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The white color of the birch bark has a protective function, because it can regulate the temperature of the tree. The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously. This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark. In addition, the betulin in the white bark protects against bacteria and fungi, as it is antibacterial and water repellent.

	Claim	Evidence	Reasoning	I don't know
The white color of the birch bark has a protective function	○	○	○	○
because it can regulate the temperature of the tree	○	○	○	○
The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously.	○	○	○	○

This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark	○	○	○	○
In addition, the betulin in the white bark protects against bacteria and fungi	○	○	○	○
as it is antibacterial and water repellent.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Claim]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Evidence]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Reasoning]

Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [I don't know]

Q25 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The white color of the birch bark has a protective function (Claim), because it can regulate the temperature of the tree (Evidence). As the low winter sun causes the bark to be exposed to direct sunlight that increases the temperature of the bark tremendously (Reasoning). This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark (Reasoning). In addition, the betulin in the white bark protects against bacteria and fungi (Evidence), as it is antibacterial and water repellent (Reasoning).

Q27
Please read the following example about a student's judgement of bee stings. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Bees die after they sting as the resulting wound on their abdomen cannot heal, this is because the sting of the bee has 10 barbs that are torn out of the abdomen during the sting. In addition, after the sting, venom continues to be pumped through the muscles into the open wound.

	Claim	Evidence	Reasoning	I don't know
Bees die after they sting	○	○	○	○
as the resulting wound on their abdomen cannot heal	○	○	○	○
this is because the sting of the bee has 10 barbs that are torn out of the abdomen during the sting.	○	○	○	○
In addition, after the sting, venom continues to be pumped through the muscles into the open wound.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Claim]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Evidence]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [Reasoning]

Or Please read the following example about a student's judgement of bee stings. After reading the ex... = In addition, after the sting, venom continues to be pumped through the muscles into the open wound. [I don't know]

Q29 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: Bees die after they sting (Claim) as the resulting wound of their abdomen cannot heal (Evidence), this is because the sting of a bee has 10 barbs that are torn out of the abdomen during the sting (Reasoning). In addition, after the sting, venom continues to be pumped through the muscles into the open wound (Reasoning).

Q31
Please read the following example about a student's judgement of the emergence of high and low tides. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The moon causes high and low tides because water moves differently depending on its position on earth's surface. This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull. Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.

	Claim	Evidence	Reasoning	I don't know
The moon causes high and low tides	○	○	○	○
because water moves differently depending on its position on earth's surface.	○	○	○	○
This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull.	○	○	○	○
Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Claim]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Evidence]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [Reasoning]

Or Please read the following example about a student's judgement of the emergence of high and low ti... = Moreover, the actually small effect of acceleration leads to high water pressure in the oceans by the collective migration of matter in large bodies of water. [I don't know]

Q33 Here you can see the solution to compare it with your choice. Please do not change your initial answer.

Solution: **The moon causes high and low tides (Claim) because water moves differently depending on its position on earth's surface (Evidence). This is because the water moves along the earth's surface towards the poles of the earth-moon-line due to the moon's pull (Reasoning). Moreover, the actually small effect of acceleration leads to a high water pressure in the oceans by the collective migration of matter in large bodies of water (Reasoning).**

Q35

Please read the following example about a student's judgement of an organic chemistry reaction step. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$, as the intermediate formed in the S_N1 reaction is stabilized, and because chloride is a good leaving group.

	Claim	Evidence	Reasoning	I don't know
In polar solvent, the S_N1 reaction of tert-butyl chloride proceeds faster with $AgNO_3$ than the S_N2 reaction of n-butyl chloride with $AgNO_3$	○	○	○	○
as the intermediate formed in the S_N1 reaction is stabilized	○	○	○	○
and because chloride is a good leaving group.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Claim]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Evidence]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [Reasoning]

Or Please read the following example about a student's judgement of an organic chemistry reaction st... = and because chloride is a good leaving group. [I don't know]

Q37 Here you can see the solution to compare it with your choice. Please do not change your initial answer.

Solution: **In a polar solvent, the S_N1 reaction of tert-butyl chloride with $AgNO_3$ proceeds faster than the S_N2 reaction of n-butyl chloride with $AgNO_3$ (Claim), as the intermediate formed in the S_N1 reaction is stabilized (Evidence), and because chloride is a good leaving group (Evidence).**

Q39 Please read the following example about a student's judgement of an electrochemistry reaction. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The reaction of zinc ions with copper is reactant favored, as copper is not able to reduce zinc ions but zinc is able to reduce copper ions, since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent.

	Claim	Evidence	Reasoning	I don't know
The reaction of zinc ions with copper is reactant favored	○	○	○	○
as copper is not able to reduce zinc ions but zinc is able to reduce copper ions	○	○	○	○
since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Claim]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Evidence]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [Reasoning]

Or Please read the following example about a student's judgement of an electrochemistry reaction. Af... = since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent. [I don't know]

Q41 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The reaction of zinc ions with copper is reactant favored (Claim), as copper is not able to reduce the zinc ions but zinc is able to reduce copper ions (Evidence), since the standard electric potential of copper is higher than that of zinc making copper more noble and copper ions the oxidizing agent (Reasoning).

Q43 Please read the following example about a student's judgement of the polarity of the bond of hydrogen fluoride. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The bond in hydrogen fluoride is polar, due to the asymmetrical electron distribution in the bond. As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.

	Claim	Evidence	Reasoning	I don't know
The bond in hydrogen fluoride is polar	○	○	○	○
due to the asymmetrical electron distribution in the bond	○	○	○	○
As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Claim]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Evidence]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Reasoning]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [I don't know]

Q45 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: **The bond in hydrogen fluoride is polar (Claim), due to the asymmetrical electron distribution in the bond (Evidence). As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen (Reasoning).**

Q47
Please read the following example about a student's judgement of ampholytes. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Water is an ampholyte since it reacts as an acid and as a base. Because oxygen has a high electronegativity and can stabilize charge well, water can act as a proton donor. Further, the lone pair of the oxygen can abstract a proton from another molecule.

	Claim	Evidence	Reasoning	I don't know
Water is an ampholyte	○	○	○	○
since it reacts as an acid and as a base	○	○	○	○
Because oxygen has a high electronegativity	○	○	○	○
and can stabilize charge well, water can act as a proton donor.	○	○	○	○
Further, the lone pair of the oxygen can abstract a proton from another molecule.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Claim]

Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Evidence]

Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Reasoning]

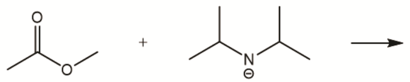
Or Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [I don't know]

Q49 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: **Water is an ampholyte (Claim), since it reacts as an acid and as a base (Evidence). Because oxygen has a high electronegativity (Reasoning) and can stabilize negative charge well, water can act as a proton donor (Reasoning). Further, the lone pair on the oxygen can abstract a proton from another molecule (Reasoning).**

Seitenumbruch

Q51 **Exercise on Conceptual Understanding**

Answer the prompts below and elaborate on your answers as detailed as possible. Please do not look up the answers on the internet or in the literature, but answers the prompts using your own knowledge only. Please do not leave any field blank. If you do not know the answer, write "I do not know" in the blank.



Answer the following tasks in as much detail as possible. In your answer do not only refer to the reactants of the reaction, but also to the products that you think form.
After typing your answers, take a photo or screenshot as you will not be able to see your answers anymore later in the survey, but you will still need them to answer further questions.

○ Decide whether steric aspects need to be considered in the reaction and explain why you think so.

____

○ Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule x has the highest and molecule y has the lowest pK_a)

____

○ Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.

____

○ Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.

____

◯ Determine whether you think there are any effects that stabilize your product compared to your reactants. If so, explain how the effect/s stabilize the product.

____

◯ Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.

____

◯ Decide whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.

____

◯ Decide whether the reaction is reactant- or product-favored from an energetic perspective (enthalpy). Explain your thinking.

____

Q53 Did you take a picture or screenshot of your answers?

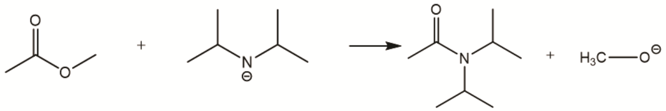
◯ Yes

◯ No

Seitenumbruch

Q55

Building arguments 1/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q103

In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.

Nucleophilicity/Electrophilicity

Nucleophilic centers: lone pairs at nitrogen and oxygen
Most electrophilic center: partial positively charged carbonyl C

Acidity/Basicity

Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.
Base: In terms of Lewis capable of donating an electron pair. Most basic atom is the nitrogen in the reactant.
Acid-base-reactions are usually faster than other reaction types.

Steric Aspects

The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.

Entropic Aspects

Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.

Electronic Aspects

Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.

Stability

A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values

CH3OH approx. 15, ester approx. 25, diisopropylamine approx. 40

Energetic Aspects

Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q57

The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

◯ The reaction product is plausible

◯ The reaction product is implausible

Q59 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q61 Answer the following question according to your personal experience.

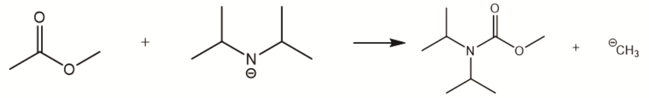
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	◯	◯	◯	◯

Q63 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	◯	◯	◯	◯

Seitenumbruch

Q65
Building arguments 2/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q105
In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.
Nucleophilicity/Electrophilicity
Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon
Most electrophilic center: partial positively charged carbonyl C
Acidity/Basicity
Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.
Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant and the negatively charged carbon in the product.
Acid-base-reactions are usually faster than other reaction types.

Steric Aspects
The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.
Entropic Aspects
Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.
Electronic Aspects
Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.
Stability
A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.
pK_a Values
$CH4$ approx. 56, ester approx. 25, diisopropylamine approx. 40
Energetic Aspects
Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q67

The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

- ○ The reaction product is plausible
- ○ The reaction product is implausible

Q69 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q71 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q73 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q75
Building arguments 3/4

Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q107
In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.
Nucleophilicity/Electrophilicity
Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon
Most electrophilic center: partial positively charged carbonyl C
Acidity/Basicity
Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group.
Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant.
Acid-base-reactions are usually faster than other reaction types.
Steric Aspects
The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.
Entropic Aspects
Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.
Electronic Aspects
Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2
Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.
Resonance effects: Distribution of electrons over several atoms. Enolates show resonance.
Stability
A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values
ester approx. 25, diisopropylamine approx. 40
Energetic Aspects
Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q77
The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q79 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		
Argument 2		
Argument 3		
Argument 4		
Argument 5		

Argument 6		

Q81 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

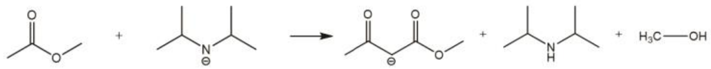
Q83 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q85

Building arguments 4/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q109

In case you had problems answering the conceptual questions, here are information you can consider for building the arguments.

Nucleophilicity/Electrophilicity

Nucleophilic centers: lone pairs at nitrogen, oxygen, and carbon
Most electrophilic center: partial positively charged carbonyl C

Acidity/Basicity

Acid: In terms of Brønsted capable of donating a proton. Most acidic proton is at carbon atom attached to the carbonyl group and in the further reaction the proton at the carbon atom between the two carbonyl groups.

Base: In terms of Lewis capable of donating an electron pair. Basic atoms are the nitrogen in the reactant and the negatively charged oxygen of methoxide as an intermediate.

Acid-base-reactions are usually faster than other reaction types.

Steric Aspects

The angle of attack at the carbonyl C is 107° (Bürgi-Dunitz angle). Bulky substituents are sterically hindered.

Entropic Aspects

Entropy indicates the distribution of energy, *e.g.*, if one reagent reacts to two products, the reaction is entropically favored.

Electronic Aspects

Electronegativity(EN): EN Oxygen = 3,5 - EN Nitrogen = 3,0 - EN Carbon = 2,5 - EN Hydrogen = 2,2

Inductive Effects: Asymmetric distribution of electrons in a bond. More electronegative atoms withdraw electrons, more electropositive atoms push electrons.

Resonance effects: Distribution of electrons over several atoms. Enolates and the product show resonance.

Stability

A molecule is stabilized if, for example, the electrons can be distributed, *e.g.*: via resonance or inductive effects.

pK_a Values

ester approx. 25, diisopropylamine approx. 40, methyl acetoacetate approx. 12

Energetic Aspects

Bond energy: C-O = 350 kJ/mol, C=O = 750 kJ/mol, C-C = 350 kJ/mol, C=C = 615 kJ/mol, C-N = 305 kJ/mol, C-H = 415 kJ/mol

Q87

The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q89 Build as many arguments as you can.

	Evidence (Support of the claim)	Reasoning (Justification of evidence)
Argument 1		

Argument 2		
Argument 3		
Argument 4		
Argument 5		
Argument 6		

Q91 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q93 Answer the following question according to your personal experience.

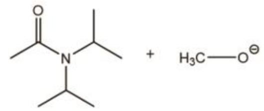
	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

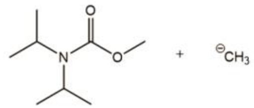
Seitenumbruch

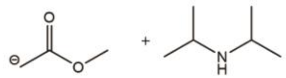
Q57 **Decision-making**

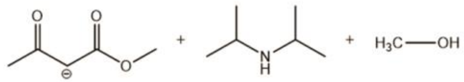
Which product is formed in the following reaction? Select the reaction product/s. Only choose final product/s of the reaction, do not select intermediate/s.

☐ I stay with the reaction product I have formed before.

☐ 

☐ 

☐ 

☐ 

Seitenumbruch

Q97 **Evaluation**

Thank you very much for completing the chemistry portion of the survey. Before you submit the survey, we would like to ask you for your feedback about how useful you thought the argumentation training was.
Please rate the following statements according to your personal experience.

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
The argumentation training helped me structure my arguments.	○	○	○	○	○
Prior to this survey, when answering tasks in organic chemistry, I rarely supported evidence with reasoning (*e.g.*, why leaving groups are good or bad)	○	○	○	○	○
The argumentation training helped me in answering the tasks.	○	○	○	○	○

I would use the argumentation pattern again when working on tasks in organic chemistry.	○	○	○	○	○
I would recommend the argumentation training to my fellow students.	○	○	○	○	○
I became more intensely involved with the tasks by building arguments.	○	○	○	○	○
I would have needed more help building arguments.	○	○	○	○	○
I was unsettled by the argumentation pattern.	○	○	○	○	○
Before the survey I did not know how arguments are constructed.	○	○	○	○	○
I would like to see building of arguments as part of the curriculum.	○	○	○	○	○
The questions about chemical concepts made me think of aspects that I had not included before.	○	○	○	○	○
The questions included key chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I found it difficult to apply these chemical concepts.	○	○	○	○	○
I was unsettled by the questions about chemical concepts.	○	○	○	○	○
I would have needed more help with questions about chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I became more intensely involved with the tasks by the questions on chemical concepts.	○	○	○	○	○

Q99 Here you have the possibility to give feedback.

____

____

____

____

____

Seitenumbruch

Q101
Thank you for participating in the second part of the study!
Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich

Ende des Blocks: Default Question Block

6.3.5 Adapted Scaffold (ReaS group) designed and used with the Software Qualtrics

Survey 2 - Group 4

Beginn des Blocks: Default Question Block

Q3 Welcome!
We (Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich) are grateful that you are part of the study.
If you have any questions, please feel free to contact Dr. Caspari (ira.caspari@tufts.edu).

Seitenumbruch

Q5 **User Code**
To maintain confidentiality in data collection, please create a code following the instructions below.

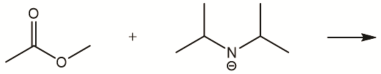
- ◯ Third letter of the street of your address (*e.g.*, 20 In**G**ram Street, Queens)
____
- ◯ Last letter of the first name of a parent/guardian (if you have two parents/guardians, choose the older one) (*e.g.*, Mar**Y** or Be**N**)
____
- ◯ Second letter of the city/town/village your parent/guardian lives in (if you have two parents/guardians, choose the older one) (*e.g.*, M**E**dford)
____
- ◯ Date of birth: Day (two digits) (*e.g.*, 08)
____

Seitenumbruch

Q7 Welcome!
In this survey you will be asked to build arguments about organic chemistry reactions. You will get specific support in how to build complete arguments and will practice this with several examples outside of organic chemistry to then apply it to organic chemistry. Please fill out this survey without any external support.

Seitenumbruch

Q9 **Task 1**
Which product is formed in the following reaction?
Draw the product *e.g.*, with Powerpoint, Word, with a drawing program like Chemdraw or take a picture of your drawing and upload the result as a file.



Q11 Build as many arguments as you can to justify your decision in as much detail as possible and in complete sentences.

- ◯ Argument 1 ________________________________
- ◯ Argument 2 ________________________________
- ◯ Argument 3 ________________________________
- ◯ Argument 4 ________________________________
- ◯ Argument 5 ________________________________

Q13 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	very unconfident
How confident are you that you have drawn the correct reaction product?	◯	◯	◯	◯

Q15 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	◯	◯	◯	◯

Seitenumbruch

Q17 **Exercise on Argumentation Patterns**
Building arguments and weighing these arguments are essential skills in science and everyday life. Often this process takes place unconsciously. For these reasons, we would like to explain and let you practice how to construct arguments in different contexts.
An argument should consist of three basic components:
The claim represents the position for which one argues. The evidence supports the claim and consists of data on which the claim is based. The reasoning justifies how the evidence relates to the claim, *e.g.*, with scientific principles.
Example: **If you want to appear taller, you should measure yourself in the morning**, **because people are on average three centimeters taller in the morning than in the evening**, **because their spinal discs lose fluid during the day and are compressed as a result.**
An argument, however, does not have to consist of just one mention of evidence or reasoning. An argument can be supported by more than one evidence and more than one line of reasoning.

Q19
Please read the following example about a student's judgement of the protective function of birch barks. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.
The white color of the birch bark has a protective function, because it can regulate the temperature of the tree. The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously. This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark. In addition, the betulin in the white bark protects against bacteria and fungi, as it is antibacterial and water repellent.

	Claim	Evidence	Reasoning	I don't know
The white color of the birch bark has a protective function	○	○	○	○
because it can regulate the temperature of the tree	○	○	○	○
The low winter sun causes the birch bark to be exposed to direct sunlight that would increase the temperature of the bark tremendously.	○	○	○	○
This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark	○	○	○	○
In addition, the betulin in the white bark protects against bacteria and fungi	○	○	○	○
as it is antibacterial and water repellent.	○	○	○	○

Diese Frage anzeigen:
If Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Claim]
Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Evidence]
Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [Reasoning]
Or Please read the following example about a student's judgement of the protective function of birch... = as it is antibacterial and water repellent. [I don't know]

Q21 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The white color of the birch bark has a protective function (Claim), **because it can regulate the temperature of the tree (Evidence). As the low winter sun causes the bark to be exposed to direct sunlight that increases the temperature of the bark tremendously (Reasoning). This increase in temperature can be reduced because the white color reflects the sun, thus, preventing cracking of the bark (Reasoning).** In addition, **the betulin in the white bark protects against bacteria and fungi (Evidence), as it is antibacterial and water repellent (Reasoning).**

Q23 Please read the following example about a student's judgement of the polarity of the bond of hydrogen fluoride. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

The bond in hydrogen fluoride is polar, due to the asymmetrical electron distribution in the bond. As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.

	Claim	Evidence	Reasoning	I don't know
The bond in hydrogen fluoride is polar	○	○	○	○
due to the asymmetrical electron distribution in the bond	○	○	○	○
As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Claim]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Evidence]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [Reasoning]

Or Please read the following example about a student's judgement of the polarity of the bond of hydr... = As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen. [I don't know]

Q25 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: The bond in hydrogen fluoride is polar (Claim), **due to the asymmetrical electron distribution in the bond (Evidence). As fluorine is more electronegative than hydrogen, it pulls the bonding electron pair closer towards itself, creating a higher electron density on the side of the fluorine and a lower electron density at the hydrogen (Reasoning).**

Q27
Please read the following example about a student's judgement of ampholytes. After reading the example, you will be asked about what parts of this example can be categorized as claim, evidence, and reasoning.

Water is an ampholyte since it reacts as an acid and as a base. Because oxygen has a high electronegativity and can stabilize charge well, water can act as a proton donor. Further, the lone pair of the oxygen can abstract a proton from another molecule.

	Claim	Evidence	Reasoning	I don't know
Water is an ampholyte	○	○	○	○
since it reacts as an acid and as a base	○	○	○	○
Because oxygen has a high electronegativity	○	○	○	○
and can stabilize charge well, water can act as a proton donor.	○	○	○	○

Further, the lone pair of the oxygen can abstract a proton from another molecule.	○	○	○	○

Diese Frage anzeigen:

If Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Claim]

And Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Evidence]

And Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [Reasoning]

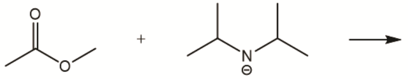
And Please read the following example about a student's judgement of ampholytes. After reading the ex... = Further, the lone pair of the oxygen can abstract a proton from another molecule. [I don't know]

Q29 Here you can see the solution to compare it with your choice. Please do not change your initial answer.
Solution: **Water is an ampholyte (Claim), since it reacts as an acid and as a base (Evidence). Because oxygen has a high electronegativity (Reasoning) and can stabilize negative charge well, water can act as a proton donor (Reasoning). Further, the lone pair on the oxygen can abstract a proton from another molecule (Reasoning).**

Seitenumbruch

Q31 **Exercise on Conceptual Understanding**

Answer the prompts below and <u>elaborate on your answers as detailed as possible</u>. Please do not look up the answers on the internet or in the literature, but answers the prompts using your own knowledge only. Please do not leave any field blank. If you do not know the answer, write "I do not know" in the blank.



Answer the following tasks in as much detail as possible. In your answer do not only refer to the reactants of the reaction, but also to the products that you think form.
After typing your answers, take a photo or screenshot as you will not be able to see your answers anymore later in the survey, but you will still need them to answer further questions.

○ Decide whether steric aspects need to be considered in the reaction and explain why you think so.

____

○ Approximate the pK_a values of the involved molecules in this reaction, or, if you do not know, outline how you think the pK_a values of the different molecules compare to each other (*e.g.*, molecule x has the highest and molecule y has the lowest pK_a)

____

○ Determine at which positions you think the involved molecules react as a nucleophile and at which positions they react as an electrophile. Explain your thinking.

____

○ Determine at which positions you think the involved molecules react as an acid and at which positions they react as a base. Explain your thinking.

____

○ Determine whether you think there are any effects that stabilize your product compared to your reactants. If so, explain how the effect/s stabilize the product.

____

○ Determine whether you think there are any entropic effects that influence the reaction process. If so, explain why you think so.

____

○ Decide whether electronic effects (*e.g.*, inductive effects, resonance, electronegativity,...) influence the reaction process and why you think so.

____

○ Decide whether the reaction is reactant- or product-favored from an energetic perspective (enthalpy). Explain your thinking.

____

Q33 Did you take a picture or screenshot of your answers?

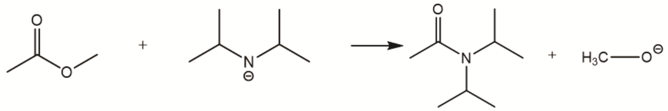
○ Yes

○ No

Seitenumbruch

Q35

Building arguments 1/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q37

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q39 Build as many arguments as you can and try to justify your argument with as many reasoning as you can.

	Evidence (Support of the claim)	**Reasoning 1** (Justification of evidence)	**Reasoning 2** (Justification of evidence)	**Reasoning 3** (Justification of evidence)
Argument 1				
Argument 2				
Argument 3				
Argument 4				

Argument 5				
Argument 6				

Q41 Answer the following question according to your personal experience.

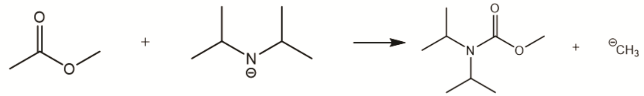
	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q43 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q45
Building arguments 2/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q47

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q49 Build as many arguments as you can and try to justify your argument with as many reasoning as you can.

	Evidence (Support of the claim)	Reasoning 1 (Justification of evidence)	Reasoning 2 (Justification of evidence)	Reasoning 3 (Justification of evidence)
Argument 1				
Argument 2				
Argument 3				
Argument 4				

Argument 5				
Argument 6				

Q51 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q53 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q55
Building arguments 3/4

Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q57

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q59 Build as many arguments as you can and try to justify your argument with as many reasoning as you can.

	Evidence (Support of the claim)	Reasoning 1 (Justification of evidence)	Reasoning 2 (Justification of evidence)	Reasoning 3 (Justification of evidence)
Argument 1				
Argument 2				
Argument 3				
Argument 4				

Argument 5				
Argument 6				

Q61 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

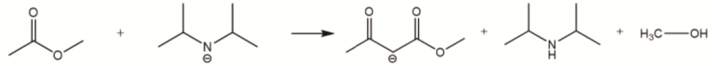
Q63 Answer the following question according to your personal experience.

	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

Seitenumbruch

Q65

Building arguments 4/4



Do you think the molecule shown is a plausible product of the reaction?
Build arguments using your answers to the conceptual questions you photographed.

Q67

Claim

○ The reaction product is plausible

○ The reaction product is implausible

Q69 Build as many arguments as you can and try to justify your argument with as many reasoning as you can.

	Evidence (Support of the claim)	**Reasoning 1** (Justification of evidence)	**Reasoning 2** (Justification of evidence)	**Reasoning 3** (Justification of evidence)
Argument 1				
Argument 2				
Argument 3				
Argument 4				

Argument 5				
Argument 6				

Q71 Answer the following question according to your personal experience.

	very confident	rather confident	rather unconfident	unconfident
How confident are you in building the arguments?	○	○	○	○

Q73 Answer the following question according to your personal experience.

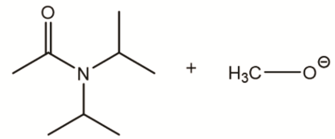
	easy	rather easy	rather difficult	difficult
How would you rate the difficulty of the task?	○	○	○	○

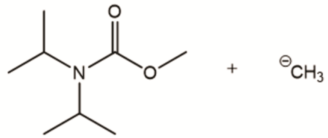
Seitenumbruch

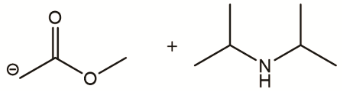
Q43 **Decision-making**

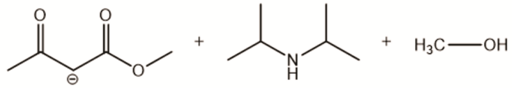
Which product is formed in the following reaction? Select the reaction product/s. Only choose final product/s of the reaction, do not select intermediate/s.

☐ I stay with the reaction product I have formed before.

☐ 

☐ 

☐ 

☐ 

Seitenumbruch

Q77 **Evaluation**

Thank you very much for completing the chemistry portion of the survey. Before you submit the survey, we would like to ask you for your feedback about how useful you thought the argumentation training was.

Please rate the following statements according to your personal experience.

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
The argumentation training helped me structure my arguments.	○	○	○	○	○
Prior to this survey, when answering tasks in organic chemistry, I rarely supported evidence with reasoning (*e.g.*, why leaving groups are good or bad)	○	○	○	○	○
The argumentation training helped me in answering the tasks.	○	○	○	○	○
I would use the argumentation pattern again when working on tasks in organic chemistry.	○	○	○	○	○
I would recommend the argumentation training to my fellow students.	○	○	○	○	○
I became more intensely involved with the tasks by building arguments.	○	○	○	○	○
I would have needed more help building arguments.	○	○	○	○	○
I was unsettled by the argumentation pattern.	○	○	○	○	○
Before the survey I did not know how arguments are constructed.	○	○	○	○	○
I would like to see building of arguments as part of the curriculum.	○	○	○	○	○
The questions about chemical concepts made me think of aspects that I had not included before.	○	○	○	○	○
The questions included key chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I found it difficult to apply these chemical concepts.	○	○	○	○	○
I was unsettled by the questions about chemical concepts.	○	○	○	○	○
I would have needed more help with questions about chemical concepts (*e.g.*, nucleophilicity).	○	○	○	○	○
I became more intensely involved with the tasks by the questions on chemical concepts.	○	○	○	○	○

Q79 Here you have the possibility to give feedback.

____

____

____

____

____

Seitenumbruch

Q81
Thank you for participating in the second part of the study!
Dr. Ira Caspari, Leonie Lieber and Dr. Nicole Graulich

Ende des Blocks: Default Question Block

6.4 Students' Original Quotes and their English Translation from Study I

In cases that the students' German utterances were grammatically incorrect, the utterance was translated into English in a grammatically correct way to facilitate reading flow. Care was always taken to ensure that the meaning of the utterance remained the same and was compared with the video recordings of the interview. The translations were checked multiple times by the coauthor and a German English student teacher. Pauses in students' utterances are indicated with […].

Table 2: Students' original quotes in German and their English translation from paper 1.

Text passage	German transcript	English translation of the transcript
Results section p. 3734	Haley: „Ja. Es war eher Automatismus, den man dann abfährt. Irgendwie halt die Reaktion und die Edukte habe ich so und so schon einmal gesehen. Das mache ich einfach. Weniger halt so gesehen dann überlegen, was, wie verhalten sie sich und so weiter und so fort, sondern eher aus der Hüfte geschossen, wie man so schön sagt."	Haley: "It was more like an automatism that I used. Somehow, I saw the reaction and reagents before. That's how I do it. I don't see it and think about how they behave and so on. It's more like 'shoot from the hip' as the phrase goes."
Results section p. 3735	Mitchell: „Und ich denke, es ist schon vorteilhaft, wenn man sich alle Möglichkeiten überlegen muss und dementsprechend auch viel mehr über die chemischen Prozesse, die im Hintergrund nachlaufen oder ablaufen, nachdenken muss."	Mitchell: "I think it is favorable to reflect on all the alternatives and therefore even more about the chemical processes that run in the background."
Results section p. 3735	Phil: „Ja. Also, ich habe das ja gemerkt, dass so die ersten beiden Aufgaben, die schaut man sich dann an, das ist automatisiert, man macht diese SN, okay, und weiter. Und sobald ich halt die anderen Produkte gesehen habe, kam dann eben dieser Zweifel und eben diese Anregungen halt zweimal darüber nachzudenken oder noch mehr. Ja, was für Möglichkeiten gibt es noch, wie sinnvoll ist es."	Phil: "I recognized that when you look at the first two subtasks, that is kind of automized. You do the SN, that's it, and you go on. As I saw the other products, I began to have doubts that made me think twice about it and even go deeper. What alternatives do I have and how plausible are they?"
Results section p. 3735	Gloria: "Das verunsichert aber auch sehr. Also man hinterfragt dann sehr viel, was man jetzt gemacht hat, und ich bin dann auch jemand, der sehr schnell dazu tendiert, den Überblick zu verlieren sozusagen."	Gloria: "These tasks made me uncertain. I questioned my decisions and I'm a person who tends to lose track of the big picture."
Results section p. 3735	Cameron: "Das Problem ist, für die Art von Aufgabe, braucht man jemanden zum Diskutieren. […] Ein Kommilitone wäre vollkommen ausreichend dafür denke ich. Es reicht ja schon, dass er sagt: "Okay, wie kommst du da drauf?" Also das ist ja schon, das	Cameron: "The problem is for these kinds of tasks that you need someone to discuss it with. […] A fellow student would be sufficient for this. It would be enough if he said 'How do you come to this conclusion?' You don't need someone who knows the correct answer, only someone asking 'Why

	heißt, man braucht gar nicht jemanden, der weiß, wie es ist, sondern man braucht einfach nur jemanden, der einen fragt: "Naja, warum denn so? Ich finde das hier sieht besser aus."	this answer?' I prefer this [approach to problem-solving]."
Results section p. 3735	Claire: „Was ich halt überhaupt nicht gemacht habe in den ersten zwei Jahre, weil da hat man was vorgesetzt bekommen, die Reaktion läuft so ab und so ist das. […] Aber eigentlich sollte ich ja denken/ mir da ein Konzept aufgebaut haben, um immer wieder darauf zurückzugreifen."	Claire: "I haven't done something like this in the last two years because you've just got something in front of you telling you that this is the mechanism of the reaction and that's it. [...] But I should have created a concept that I could refer back to."
Results section p. 3735	Claire: „Und nicht so, wie die letzten Jahre, wo dann halt, wie auch in OC1 und -2, einfach die Sachen an die Tafel geschrieben worden sind und du das dann abgeschrieben hast, dann hättest du auch von einem Lehrbuch abschreiben können."	Claire: "During [my studies in] the last few years, things have just been written in the board and you copied it down. You could've just copied the textbook instead."

Table 3: Students' original quotes in German and their English translation from paper 2.

Text passage	German transcript	English translation of the transcript
Table 1 p. 42	„Hier ist halt die Bindung schon mal stärker polarisiert."	"This bond is more polarised."
Table 1 p. 42	„Also die Bindung von CH ist halt stärker. Man muss halt mehr Energie aufwenden, um die Bindung zu brechen."	"The C-H bond is stronger. You have to spend energy to cleave the bond."
Table 1 p. 42	„… dass Säure-Basen Reaktionen sehr schnell ablaufen."	"… that acid-base reactions react very fast."
Table 1 p. 42	„… weil da halt die Gruppen einfach näher aneinander sind."	"… because the groups are close to each other."
Table 1 p. 42	„Phenole könnte man zum Beispiel mit OH^- deprotonieren."	"You can deprotonate phenols with hydroxide."
Table 1 p. 42	„… dass hier diese Stelle azider ist."	"… that this part is more acidic."
Table 1 p. 42	"Es kommt ein bisschen drauf an, wie viel Base ich zugebe."	"It depends on how much base I add."
Table 1 p. 42	„… weil das Produkt so in keiner Weise halt stabil ist"	"… this product would be definitely not stable."

Fig. 5 p. 44	Phil: „Das Argument dafür wäre, wenn nur sehr wenig Base zugesetzt wird wahrscheinlich am meisten das hier (Alkoholat). Also wenn wir jetzt 1 zu 1 Umsetzung hätten, würde wahrscheinlich das hier rauskommen, weil wir keine SN2 machen, da die Säure-Base-Reaktion viel schneller ist. Da wird eher alles quantitativ einfach nur deprotoniert. Also das kann ich auf jeden Fall nachvollziehen.“ Interviewer: „Bevor du weitergehst, du hast jetzt schon mehrfach gesagt, die Säure-Base-Reaktion ist einfach viel viel schneller als die SN in dem Fall. Warum?“ Phil: „Ja, das hat ja auch was mit der Härte zutun. Hier haben wir zum Beispiel eine harte Base und ein Proton ist eine harte Säure und nach dem HSAB Prinzip zum Beispiel ist der Energiegewinn da am besten. Ich habe einmal eine positive Ladung, also ein leeres Orbital in das ich Ladungsdichte reinschieben kann. Hier müsste man ja SN2, habe ich ja hier ein antibindendes Orbital, in das ich meine Ladung reinstecke. Ist aber halt einfach auch sterische Gründe. Es kann ja nur von da angreifen, allein das ist ja schon ein Faktor dann, weil das hier ist ja komplett endständig, es könnte ja praktisch von überall das Proton nehmen und hier muss es ja spezifisch angreifen und ja.“	Phil: “The argument for this would be, if only a very small amount of base is added, probably this (alkoxide) will be formed. If we have a 1 to 1 conversion, this would probably come out because we don’t do S_N2 because the acid-base reaction is much faster. Rather everything is simply deprotonated quantitatively. So I can understand that in any case.” Interviewer: “You already said the acid-base reaction is way faster than the S_N. Why?” Phil: “Yes, that has something to do with hardness. Here we have a hard base and a proton is a hard acid and according to the HSAB principle the energy gain is best there. I have a positive charge, so I have an empty orbital in which I can push charge density. With an S_N2 I have an antibonding orbital into which I can insert my charge. But there are also steric reasons. It can only attack from there, that alone is a factor, because this is completely terminal, it could take the proton from practically anywhere, and here it must attack specifically.”
Fig. 6 p. 45	Interviewer: „Würdest du sagen, ab Schritt 2 ist es plausibel?“ Dylan: „Der Teil ja, weil ich hier eine partial positive Ladung, ich habe eine negative Ladung und dann ich hab keine große Ringspannung. Also ein Fünfring ist auch eher stabil und Chlorid ist eine gute Abgangsgruppe.“ […] Interviewer: „Könntest du mir probieren zu erklären, warum Chlorid also was die Erklärung dafür ist, dass Chlorid eine gute Abgangsgruppe ist.“	Interviewer: “Do you think this step is plausible?” Dylan: “Yes, because here is a partial positive charge, there is a negative charge. There is also just a little ring strain. A five-membered ring is quite stable and chloride is a good leaving group.” […] Interviewer: “Could you try to explain why chloride is a good leaving group?” Dylan: “Chloride has an electron configuration and a high electronegativity. Compared to carbon, chloride is more

	Dylan: „Ja es hat quasi Edelgaskonfiguration. Und, ja, hat auch eine hohe Elektronegativität, jetzt mit einem Bindungspartner jetzt als Kohlenstoff zum Beispiel, ehm, ist es eh partial negativ geladen. […] Thermodynamisch günstiger, jetzt so. dass ist jetzt also die Bindung ist für das Chlorid günstiger als Ionen vorzuliegen.“	electronegative. […] As an ion, chloride is thermodynamically more stable.”
Fig. 7 p. 45	Interviewer: “Jetzt hast du gerade über gute und schlechte Abgangsgruppen gesprochen. Was ist denn für oder warum ist Chlor für dich eine bessere Abgangsgruppe als OH^-.“ Amber: „Weiß ich nicht, keine Ahnung (lacht).“ Interviewer: „Das heißt, dass ist wieder ein Fakt, den du halt weißt?“ Amber: „Genau, den habe ich irgendwann mal gelernt und das quasi entweder nicht hinterfragt oder ja, einfach so hingenommen.“ Interviewer: „Könntest du probieren anhand von Chlor und OH- zu erklären, warum Chlor besser rausgeht?“ Amber: „Vielleicht weil Chlorid kleiner ist als OH-, aber wissen tu ich es nicht.“	Interviewer: “You have talked about good and bad leaving groups. Why do you think that chloride is a better leaving group than hydroxide?” Amber: “I have no idea (laughing).” Interviewer: “Is that a fact you just know?” Amber: “Exactly, I learned that at some point and either didn’t question it or simply accepted it.” Interviewer: “Could you try to explain it?” Amber: “Maybe because chloride is smaller than hydroxide but I don’t know it.”
Results section p. 46	Charlie: „Ich halte es nicht für nicht plausibel. Ich wäre jetzt intuitiv nicht drauf gekommen, dass sowas passieren würde, wenn ich ehrlich bin. Dafür hätte ich halt gleich den simpelsten Weg gegangen und hätte gesagt es ist eine ganz normale nukleophile Substitution halt eben an der Stelle, wenn ich mir das so angucke würde ich jetzt nicht unbedingt sagen, es könnte ein Nebenprodukt sein, was entsteht. Aber das Hauptprodukt wäre halt, je mehr ich darüber nachdenke desto faszinierender finde ich das.“ Interviewer: „Warum?“	Charlie: “‘‘I don’t think it’s not plausible (he refers to THF). I wouldn’t have intuitively guessed that something like this (THF) would happen if I am honest. But I would have taken an easier way and said it is a normal nucleophilic substitution, at this point, if I look at it that way I would not necessarily say it (THF) could be a by-product of what is being created. But the main product, the more I think about it, the more fascinating I think it is.’’ Interviewer: ‘‘Why?’’ Charlie: ‘‘Because I really didn’t think that this (formation of THF) was a possibility. Well, I didn’t think about it at all, I even

	Charlie: „Weil ich wirklich nicht darüber nachgedacht habe, dass das eine Möglichkeit ist. Also ich habe absolut nicht daran gedacht, ich habe bei der Aufgabe vorher, ich glaube sogar, dass das falsch ist, was da vorher war, weil ich habe halt gesehen, ich habe ein Nukleophil und eine gute Abgangsgruppe und zack nukleophile Substitution, aber ich habe nicht gesehen, dass man ja quasi sich ein weiteres Nucleophil bauen kann, wenn man so will.“	think that what was there before (nucleophilic substitution) is wrong, because I just saw that I have a nucleophile and a good leaving group and nucleophilic substitution, but I didn't see that you can build another nucleophile (alkoxide).''
Results section p. 46	Interviewer: „Wir können es ja mal genau andersherum machen. Wenn du das Produkt siehst, würdest du das Produkt denn prinzipiell als plausibel bezeichnen?“ Sonia: „Ja, aber ich weiß nicht, warum. Das sagt mir wieder mein Gefühl, ich weiß nicht. Das sieht irgendwie so richtig aus.“ Interviewer: „Und gibt es irgendwelche Faktoren, an denen du festmachen würdest, das ist richtig oder deshalb denke ich, dass das Produkt richtig sein könnte?“ Sonia: „Wir hatten mal so eine ähnliche Aufgabe in der Übung und daran ich weiß nicht, als ich das gesehen habe, hat es irgendwie klick gemacht in meinem Kopf, irgendwo in der hintersten Ecke, wo ich mir gedacht habe, ich glaube, das ist es.“	Interviewer: ''We can do it differently. When you see the product, would you describe the product as plausible in principle?'' Sonia: ''Yes, but I don't know why. My feeling tells me that again, I don't know. It looks so right somehow.'' Interviewer: ''And are there any factors that help you determine why you think the product could be right?'' Sonia: ''We once had a similar task in an exercise and I don't know, when I saw that, it kind of clicked in my head, somewhere in the back corner where I thought, I think that's it.''
Results section p. 46	Andy: „Also die OH-Gruppe wurde einfach deprotoniert.“ Interviewer: „Für wie plausibel hältst du das?“ Andy: „Das ist eine gute Frage. Eigentlich, also es kann sein, dass es auch passiert. Aber negative Ladung nur am Sauerstoff ist sehr, ich würde sagen, unwahrscheinlich.“ Interviewer: „Warum?“ […]	Andy: ''So here, the hydroxyl group was simply deprotonated.'' Interviewer: ''How plausible do you think is this?'' Andy: ''That's a good question. Actually, so it may happen. But a negative charge only on oxygen is very, I would say, unlikely.'' Interviewer: ''Why?'' […] Andy: ''So, that charge is on the oxygen and then is not somehow stabilised. That is

	Andy: „Also ich finde den ein bisschen/ Ja, dass Ladung irgendwie am Sauerstoff ist und dann halt nicht irgendwie stabilisiert wird, finde ich ein bisschen seltsam. Weil, Ladung an einer Estergruppe wäre ja irgendwie, finde ich, normal. (...) Also bei einer Estergruppe hätte man ja/ Dass da Ladung da ist, würde ich normal finden, weil das ja relativ stabilisiert ist durch hier Mesomerie. Aber nur am O, ich finde es ein bisschen seltsam, würde ich sagen.“	a bit strange to me. Because, charge on an ester group would be kind of, I think, normal, because that’s relatively stabilised by resonance. But only at the oxygen is a bit strange to me.’’
Results section p. 47	Interviewer: „Für wie plausibel hältst du das?“ Andy: „Das könnte sogar plausibler sein als das hier.“ Interviewer: „Okay. Warum?“ Andy: „Ja, das O hat freie Elektronenpaare, ein Fünfring ist nicht so gespannt wie ein Vierring zum Beispiel. Also dass es hier angreift, ist halt/ Ein Vierring ist halt ziemlich gespannt, Fünfring ja eher nicht so. Das könnte möglich sein, würde ich sagen. Also ich würde es sogar als relativ plausibel einschätzen.“	Interviewer: ‘’How plausible do you think is it?’’ Andy: ‘’That might even be more plausible than this (diol).’’ Interviewer: ‘’Okay. Why?’’ Andy: ‘’The O has lone pairs, a five-membered ring is not as strained as a four-membered-ring, for example. I would say that could be possible. So, I would even consider it relatively plausible.’’
Results section p. 47	Andy: “Das O hat, wie gesagt, freie Elektronenpaare, ist mir dann aufgefallen. Und habe dann gedacht, ja, um den Fünfring zu rekonstruieren, muss es ja hier angreifen. Da habe ich ja erst das Cl hingemalt. Es wäre aber ein fünfbindiger Kohlenstoff, der es dann ja nicht gibt. Also muss das Cl auch irgendwie rausfliegen. Und im letzten Schritt halt wird das H hier vom OH angenommen. Aber auch nur, weil es/ Genau, weil das O halt positiv ist und es möchte ja relativ neutral werden. Und es muss ja irgendwie das Furan oder auch, wenn es nicht das Furan ist, rauskommen.“ Interviewer: „Das heißt, du hast im Prinzip geguckt, ich kenne das Produkt. Wie wurschtel ich es zusammen.“	Andy: ‘’I noticed that the O has lone pairs. And then I thought, in order to reconstruct the five-membered-ring, it must attack here. That’s where I first drew the Cl. But it would be a pentavalent carbon atom, which does not exist. So, the Cl has to get out somehow. And in the last step the H is taken by the OH. But only because the O is positive and it wants to become relatively neutral. And somehow furan (Andy confounded THF and furan) has to be the product.’’ Interviewer: ‘’That means, you basically made sure that you knew the product and how to put it together?’’ Andy: ‘’Yes, somehow. Exactly.’’

	Andy: „Ja, irgendwie. Genau“	
Results section p. 47	“Wegen dem +I-Effekt, das ist ja ein elektronenschiebender Effekt, würden zu einem negativen Punkt noch mal Elektronen hingeschoben werden“ (Ronnie)	‘‘due to the positive inductive effect, which is an electron-pushing effect, more electrons would be distributed’’ (Ronnie)
Results section p. 47	„negative Ladungen neben negativer Ladung ist eher ungünstig, weil dann viel zu vieler Elektronenüberschuss ist.“ (Haley)	‘‘negative charge next to negative charge is unfavourable because there is a high excess of electrons’’ (Haley)
Results section p. 47	„Weil die Entropie hier steigt. Das hat wieder was mit der Entartung der Energie zu tun. Weil umso mehr Moleküle, also wenn die Energie auf mehr Moleküle verteilt wird, ist es am Ende stabiler also energetisch günstiger.“ (Pepper)	‘‘that is because the entropy increases. The degeneracy in the system increases, since the energy can be distributed over more molecules.’’ (Pepper)
Results section p. 48	„Weil es als Ion irgendwie halt kleiner ist und ein bisschen beweglicher.“ (Stella)	‘‘that chloride, as an ion, is smaller and can move easily compared to the hydroxide ion.’’ (Stella)
Results section p. 48	„Weil die Kette sage ich jetzt mal ziemlich lang ist.“ (Sal)	‘‘because the carbon chain is quite long.’’ (Sal)
Results section p. 48	Cameron: „Chlorid ist eine gute Abgangsgruppe, weil, ich denke, dass die Chlorkohlenstoffbindung nicht besonders stabil/ also schon stabil, aber nicht überragend im Gegensatz zu Abgangsgruppen, wie zum Beispiel OH oder OH- als Abgangsgruppe. Dann kann es in wässriger Umgebung gut stabilisiert werden. […] Es ist halt auch eine hohe Elektronegativität. Das ist auch sehr elektronenziehend, Chlor als Atom.“ Interviewer: „Könntest du erklären, warum Chlor eine gute und Hydroxid eine schlechte Abgangsgruppe ist?“ Cameron: „[…] Also theoretisch müsste es so sein, dass das Elektron, was beim Chlor hinzukommt, in ein Orbital gelagert wird, was elektronisch höher ist als beim OH-/ Energetisch, nicht elektronisch, energetischer ist als beim OH-. Aber so richtig weiter komme ich damit nicht. (lacht) Also ich würde jetzt sagen, so, dass das gar nicht aufgrund der Stabilität der	Cameron: ‘‘Chloride is a good leaving group because the chlorocarbon bond is not very stable, unlike leaving groups such as OH-. Then it can be stabilised in an aqueous solution. [. . .] There is also a high electronegativity. It’s also very electron-withdrawing, chlorine as an atom.’’ Interviewer: ‘‘Could you explain to me why chloride is a good leaving group and hydroxide is a weak leaving group?’’ Cameron: ‘‘[. . .] So theoretically it should be that the electron that is added in chlorine is stored in an orbital, which is energetically higher than hydroxide. But I’m not really getting anywhere with it (laughs). So, I would say that this is not the cause because of the stability of the two ions, but because of the stability of the bond. That the splitting, here with chlorine and then there is another bond with carbon, that the splitting, if I were to occupy that as well, of the orbitals is lower with chlorine than with oxygen. And therefore, if this energy gain, through the bond or the energy loss through the

	beiden Ionen der Fall ist, sondern aufgrund der Stabilität der Bindung, dass die Aufspaltung, wenn man sich das anguckt, hier mit Chlor und dann gibt es noch eine Bindung mit Kohlenstoff, dass dann hier die Aufspaltung, wenn ich das jetzt auch noch besetzen würde, von den Orbitalen beim Chlor geringer ist als beim Sauerstoff. Und deswegen, wenn dieser Energiegewinn, durch die Bindung beziehungsweise der Energieverlust durch das Brechen der Bindung, dass der dann geringer ist, beim Chlor als beim OH- bzw. als beim Hydroxidion."	cleavage of the bond, is then lower with chlorine than with OH.''
Results section p. 48	Gloria: "Halogene sind allgemein eine gute Abgangsgruppe. Deshalb vermute ich eine SN Reaktion." Interviewer: „Kannst du mir begründen, wieso Halogene gute Abgangsgruppen sind?" Gloria: „Weil speziell Cl- relativ stabil ist und die Oktettregel erfüllt."	Gloria: ''Halogens are generally a good leaving group. That's why I suspect a nucleophilic substitution.'' Interviewer: ''Can you tell me why halogens are good leaving groups?'' Gloria: ''Because Cl- in particular is relatively stable and obeys the octet rule.''

7 Acknowledgements

Zum Abschluss möchte ich noch den Menschen danken, ohne deren Unterstützung diese Promotion nicht möglich gewesen wäre.

Meiner Doktormutter Prof. Dr. Nicole Graulich

Liebe Nicole, dir muss und möchte ich für alles danken, was auf diesem Weg passiert ist. Ohne dich wäre einfach nichts von dieser Arbeit möglich gewesen. Du hast mich immer unterstützt, egal zu welcher Tageszeit. Du gibst wirklich alles für uns. Ich bin dankbar, dass ich mit allem zu dir kommen kann. Wir können stundenlang miteinander diskutieren und zu neuen Ideen kommen. Besonders dankbar bin ich dir, dass ich auch anderer Meinung sein durfte. Wir konnten immer konstruktiv miteinander sprechen und ich wusste, dass du immer eine offene Tür und ein offenes Ohr hast, ganz gleich was es betrifft. Du wusstest genau, wie du mich fordern und fördern kannst und auch wenn ich geflucht habe, wusste du am Ende oft besser, was das Richtige für mich ist. Du hast probiert alles möglich zu machen, Türen zu öffnen und mir den Weg als Nachwuchswissenschaftlerin geebnet. Du bist für mich ein großes Vorbild, wie du alles rockst und wie besonnen du bleibst. Neben der beruflichen Zusammenarbeit, habe ich auch jedes Gespräch in der Teeküche und jede gemeinsame Aktivität, egal ob bei euch im Hof oder bei AG-Ausflügen genossen. Ich kann dir wirklich nicht oft genug sagen, dass du eine super Chefin warst und ich froh bin, in deiner Arbeitsgruppe geforscht zu haben.

Prof. Dr. Ira Caspari-Gnann

Liebe Ira, du bist der Stein, der alles ins Rollen gebracht hat. Bereits als ich dein HiWi sein durfte und als du mich bei meiner wissenschaftlichen Hausarbeit betreut hast, warst du ein großes Vorbild für mich. Wegen dir habe ich die Promotion in Erwägung gezogen und dank der unzähligen gemeinsamen Gespräche und Momente hast du mich immer weiter bestärkt. Ich schaue zu dir auf und bin froh, dass du Teil meines Weges warst. Du hast mir Türen geöffnet und dafür werde ich dir nie genug danken können. Als dann die Möglichkeit bestanden hat, eine Kooperation mit dir an der Tufts University in Boston starten zu können, war ich Feuer und Flamme. Auch hier war die Zusammenarbeit fast spielerisch. Ich bin wahnsinnig froh, dass du Teil dieses Weges warst und hoffe, dass der Weg noch nicht zu Ende ist.

Prof. Dr. Michael Schween und PD Dr. Sascha Bernholt

Lieber Michael, lieber Sascha, ich danke euch für euer stets offenes Ohr. Ich konnte mich immer bei euch melden und ihr standet mir mit Rat und Tat zu Seite, um meine Promotion weiterzubringen. Ihr habt Lösungsvorschläge geliefert und mir im Aufgabendesign und beim Analysieren der Daten wertvolle Hinweise geliefert, vielen Dank dafür.

Allen Mitgliedern der Prüfungskommission: Prof. Dr. Nicole Graulich, Prof. Dr. Richard Göttlich, Prof. Dr. Siegfried Schindler, Prof. Dr. Michael Schween

Liebe Prüfungskommission, ihr seid die letzte Hürde dieser Promotion. Ohne euren kritischen Blick, ohne eure kritischen Fragen hätte ich nicht die Chance meine Arbeit verteidigen zu dürfen. Ich danke euch für die Zeit, die ihr euch meinetwegen nehmt.

Allen ehemaligen und gegenwärtigen Kolleginnen und Kollegen des Instituts für Didaktik der Chemie: Prof. Dr. Nicole Graulich, Susanne Luch, Nina Clemens, Dorothea Stasch, Claudia Neidel, Dr. Annette Geuther, Heiko Barth, Bettina Romberg, Silke Wagner, Dr. Stefanie Lenzer, Dr. Liz Keiner, Julia Eckhard, Julia Ortmann, Irina Braun, Axel Langner, Lukas Groos, David Kranz und Paul Martin.

Liebe AG, bereits in meinem Studium durfte ich einen Großteil von euch kennenlernen und hatte schon als Studierende das Gefühl, gut aufgehoben zu sein. In meiner Promotion konnte ich immer auf euch zählen, eure Unterstützung erhalten, viele witzige, aber auch tiefgründige Gespräche führen. Ich bin immer gern zur Arbeit gekommen und habe mich auf euch gefreut, das ist nicht selbstverständlich. Ihr seid herzliche, liebevolle und humorvolle Kolleginnen und Kollegen gewesen und legt die Messlatte für alles, was danach kommt, wahnsinnig hoch. Ihr seid alle fantastisch und ihr werdet mir sehr fehlen!

Liebe Frau Geuther, Sie haben mich bereits im Studium geprägt und ich werde viele Aspekte Ihrer Lehre nie vergessen. Sie haben immer wieder neue Perspektiven aufgezeigt und durch Ihre Organisation dazu beigetragen, dass alles reibungslos ablaufen konnte. Ihre Ratschläge und die Gespräche mit Ihnen habe ich immer als sehr wertvoll aufgefasst.

Liebe Silke, du bringst jeden zum Strahlen, wenn du den Raum betrittst. Deine fröhliche, offene Art zieht jeden in ihren Bann und sorgt dafür, dass wir stundenlange Gespräche führen konnten. Du hast Diskussionen initiiert, die zum Nachdenken angeregt haben und stellst für mich die Lehrerin dar, die ich einmal sein möchte. Du lebst für deinen Beruf, der wohl eher eine Berufung ist, was mich immer wieder beeindruckt. Ich danke dir für jeden Impuls, jedes Gespräch und jedes gemeinsame Lachen.

Lieber Heiko, jeder Tag, den ich mit dir verbringen durfte, war ein guter Tag. Du bist das Herz der AG, völlig zurecht. Ich durfte sehr viel Zeit mit dir verbringen, unzählige Gespräche mit dir führen. Du hast mir immer mit Rat und Tat zur Seite gestanden, wir konnten uns über Gott und die Welt unterhalten, uns zusammen aufregen und vor allem viel zusammen lachen. Du sagst mir die Meinung, bist ehrlich zu mir. Das ist wohl die Eigenschaft, die ich am meisten an dir schätze. Die gemeinsamen Autofahrten ins „schlechte“ Buseck. Geheimnisse, die wir einander anvertraut haben. Du bist liebevoll und impulsiv, suchst Nähe und brauchst Abstand. Du bist einfach der beste Kollege, den ich mir hätte vorstellen können. Allein beim Schreiben dieser Zeilen steigen mir schon die Tränen in den Augen. Ich werde unser Zuwinken vermissen, die Gespräche über veganes Essen und Hackfleischpizza, über Politik und die Formel 1 vermissen. Ich werde einfach dich vermissen. Du bist ein fantastischer Mensch, vergiss das nie. Danke, dass ich Teil deines Lebens sein darf.

Liebe Liz, Frau Doktor, ich glaube ich habe selten so gelacht wie mit dir. Wir kennen uns seit mittlerweile neun Jahren, doch erst die Promotion hat uns so richtig zusammengeführt. Wir konnten unzählige Gespräche miteinander führen, uns über die gleichen Dinge aufregen. Wir haben uns immer unterstützt, uns füreinander gefreut und einander ausgelacht. In manchen Dingen sind wir uns so ähnlich und in manchen so unterschiedlich, das hat es ausgemacht. Wir konnten miteinander diskutieren, aber vor allem einander helfen. Wir konnten uns immer aufeinander verlassen und die Welt war ein bisschen weniger leuchtend als du gegangen bist. Neue Lebensabschnitte beginnen heißt aber nicht, dass man sich auseinanderleben muss. Diese Hoffnung hat sich bestätigt und ich bin wirklich froh, dass du Teil meines Weges warst und bist. Du bist spitzenklasse, danke!

Liebe Jule, du hast mich durch die ganze Promotion begleitet und warst immer ein Vorbild für mich. Du bist für mich die Verkörperung einer Wissenschaftlerin, du bist ehrgeizig, strauchelst und stehst stärker wieder auf. Ich bewundere dich und glaube, dass dir eine große Zukunft bevorsteht. Ich danke dir für all die schönen Gespräche, viele ernste, aber noch mehr lustige. Ich konnte viel von dir lernen, du hast mich unterstützt und ich bin froh, dass auch ich dir hin und wieder eine Hilfe sein konnte. Ich wünschte unsere Promotionen wären nicht von so vielen externen Faktoren beeinflusst worden und wir hätten gemeinsam nach Edinburgh und an viele weitere tolle Orte reisen können. Wir beschreiten gleichzeitig den nächsten Schritt und ich hoffe, dass sich unsere Wege dabei oft kreuzen werden. Vergiss nie, wir sind fabelhaft!

Liebe Julia, wir haben uns bereits vor deiner Zeit als Doktorandin kennengelernt und viel über deine Pläne gesprochen. Deine immer fröhliche und herzliche Art bringt mich selbst immer wieder zum Lachen. Du bist ein sehr offener Mensch, gehst auf dein Umfeld zu und stellst Fragen. Durch deine Art habe ich mich immer wohl in deiner Nähe gefühlt und durch deine Reiseerfahrung eine eigene Bucketlist erstellt. Ich hoffe, dass du dir deine lebensfrohe Art stets beibehalten wirst und danke dir, dass du mich immer wieder zum Lachen gebracht hast.

Liebe Irina, deine rationale Denkweise hat häufig dazu geführt, dass ich Dinge noch einmal überdacht habe. Du bist immer zuverlässig und ich konnte dir blind vertrauen, wenn du Arbeit übernommen hast. Vor allem haben wir aber auch viel gelacht, Filmrollen besetzt und uns über gutes Essen unterhalten. Außerdem warst du meine persönliche Sprachlehrerin, Спасибо und merci.

Lieber Axel, wo soll ich da anfangen? Unsere Beziehung war von 0 auf 100. Zugegeben, für die 0 kannst du nichts, das hast du jemand anderem zu verdanken. Deine humorvolle Art ist unschlagbar, oft provokativ, aber da geben wir uns nichts. Ich habe in so kurzer Zeit noch nie so viel mit einem Menschen gelacht. Wir können über alles miteinander sprechen, wirklich alles. Wir verstellen uns nicht, sondern sind einfach ehrlich zueinander. Das macht dich und unsere Freundschaft aus. Ich kann mich immer auf dich verlassen, du bist immer da. Wir können stundenlang reden, ganz allgemein, aber vor allem sehr privat. Du kennst meine Art und nimmst sie an. Ich vertraue dir, bin dankbar für deine Hilfe und deine Meinung. Unsere Diskussionen können sich in die Länge ziehen, wir sind nicht immer einer Meinung, aber genau dafür sind Freunde da. Ich genieße es Zeit mit dir zu verbringen, ob privat oder an der Arbeit und ich freue mich jeden Tag, dir als frühem Vogel als erstes guten Morgen zu sagen. Du beeindruckst mich mit deiner Arbeitshaltung, immer mit vollem Herzblut dabei. Dank dir ist Kleonie Klieber zu einem besseren Menschen geworden, ich kann nicht oft genug danke sagen.

Lieber David, wir gehen seit neun Jahren durch dick und dünn. Wir sind fast wie ein altes Ehepaar. Ich kann mich immer auf dich verlassen, vor allem auf tägliche Anrufe, hör bloß nicht auf damit. Wir sind unsere gegenseitigen Kummerkästen, erzählen uns praktisch alles. Bei dir fühle ich mich sicher. Ich kann dich jederzeit um Rat zu fragen, zu jeder Tages- und Nachtzeit. Wir waren nie Kollegen, sondern immer Freunde und ich bin mir sicher, dass wir das auch noch viele Jahre sind. Unsere Lebensstile sind zwar sehr unterschiedlich, aber trotzdem verstehen wir uns blind. Ich habe dich einfach super gern und bin froh, dass ich diesen Weg mit dir gemeinsam bestreiten durfte.

Lieber Paul, wir haben uns während deines Studiums kennengelernt und schnell gemerkt, dass wir auch außerhalb des universitären Umfelds wirklich gut miteinander harmonieren. Wir können immer gute und vor allem produktive und konstruktive Gespräche führen. Es macht mich sehr stolz, dass du an meinem Promotionsprojekt, meinem Baby, weiterarbeiten wirst. Wir haben uns vor allem in deinen ersten und meinen letzten Wochen viel unterstützt, gegenseitig Hilfestellungen gegeben und großartige Themen diskutiert. Ich kann mich dabei immer auf dich verlassen und danke dir auch für den Beitrag, den du zu dieser Dissertation beigetragen hast. In dir steckt wahnsinnig viel Potential und ich freue mich, deinen weiteren Weg weiter verfolgen zu können.

Lieber Lars, du bist das inoffizielle Mitglied der AG, für mich gehörst du aber fest dazu. Wir bestreiten das universitäre Leben seit fast 10 Jahren zusammen und waren lange Zeit fester Bestandteil voneinander. Ohne dich hätte das Studium nicht mal halb so viel Spaß gemacht, ohne dich hätte ich nie den Bachelor gemacht und wahrscheinlich die Promotion auch nicht in diesem Maße angestrebt. Auch wenn uns mittlerweile zwei Stockwerke voneinander trennen, freue ich mich über jede gemeinsame Kaffeepause. Ich kann mir arbeiten ohne dich nicht vorstellen, will es auch nicht. Allein beim Gedanken daran, dass sich unsere Wege bald beruflich trennen werden blutet mein Herz. Über all die Jahre hat sich aber so eine enge Freundschaft entwickelt, dass ich mir sicher bin, dass wir privat weiterhin ein Team sind. Ich hätte mir keinen besseren Lernpartner wünschen können und bin dankbar für jede Diskussion, jedes Gespräch, jedes Gefluche und jeden Lacher. Ich bin dir einfach dankbar.

allen Studierenden, die mich als Probandinnen und Probanden, in der Forschung oder in der Lehre begleitet haben

Liebe Studierende, ohne eure Hilfe und Beteiligung wäre diese Arbeit nicht möglich gewesen. Ihr habt mir Daten geliefert, habt eure Zeit geopfert, habt mir beim Auswerten geholfen oder mir Arbeit im Labor abgenommen. Ihr habt die vielen Stunden Arbeit erleichtert, vielen Dank dafür.

Liebe Krenare, du bist seit fast der ersten Stunde mein HiWi. Für diesen Job hätte ich mir niemand besseren vorstellen können. Du bist von Tag zu Tag immer stärker mit meiner Forschung verwurzelt und ohne dich wäre vieles schwieriger gewesen. Ich bin stolz, dass wir sogar gemeinsam publiziert haben. Du bist in der ganzen Zeit viel mehr als ein HiWi, sondern eine Freundin gewesen. Du hast mich gepusht, warst mein Fan. Wir konnten über so viel reden, uns anvertrauen und so viel lachen. Dein Humor ist einzigartig, behalte ihn dir bei. Ich danke dir, dass du mich auf meinem Weg begleitet hast.

meiner Familie

Liebe Mama, lieber Papa, ihr habt mich einfach immer unterstützt, seitdem ich klein bin. Auch wenn ihr nicht immer versteht, was ich bei meiner Arbeit eigentlich getan habe, gebt ihr euch alle Mühe. Ihr seht nichts als selbstverständlich an, sondern wisst wie viel Arbeit in allem steckt. Euch habe ich alles zu verdanken, eure grenzenlose Liebe hat immer dafür gesorgt, dass ich mich sicher fühle. Ich kann mit jedem Problem zu euch kommen und ich weiß, dass wir für alles eine Lösung finden. Ich weiß wie unfassbar stolz ihr auf mich seid und ich könnte nicht stolzer sein, euch als meine Eltern zu haben. Ihr habt mich geprägt wie niemand anderes, mir vorgelebt wie es ist glücklich zu sein. Ihr ermutigt mich bei jedem Schritt, steht hinter mir, egal was ich tue oder für was ich mich entscheide. Ich möchte euch einfach für alles danken, insbesondere für euren Respekt, dafür, dass ihr mir alles geboten habt, was euch möglich war und dass Jana und ich immer an erster Stelle standen. Ich liebe euch einfach, das wird sie nie ändern.

Liebe Jana, meine kleine Schwester, mit niemanden habe ich mehr gemeinsam als mit dir. Ich kann alles mit dir besprechen, kann Szenarien diskutieren und mit dir meinem Arbeitsalltag entfliehen. Gemeinsame Zeit ist das wertvollste und ich genieße jeden Spaziergang zu Paul und jeden Tag, den wir gemeinsam verbringen. Ich habe dich mein ganzes Leben beschützt und weiß, wie sicher ich auch bei dir bin. Kein Band ist enger als das zwischen Geschwistern, trotzdem ist unsere Beziehung mehr als besonders. Du hast deinen eigenen Willen, erkundest die Welt und egal wo du dich aufhältst, ich bin immer deine große Schwester, die auf dich aufpasst und dir immer ein Zuhause bieten wird. Ich lerne jeden Tag von dir, du bist fantastisch und meine Vertraute. Ich liebe dich.

Lieber Michi, mein Schatz, seit Tag 1 begleitest du mich in der Promotion. Du warst und bist immer mein Fels in der Brandung. Du unterstützt mich jeden Tag und zeigst mir wie stolz du auf mich bist. Du diskutierst jedes Problem mit mir, regst dich mit mir auf und freust dich für mich. Ich freue mich jeden Tag auf die Arbeit, aber noch mehr freue ich mich jeden Tag zu dir nachhause zu kommen. Du bist immer an meiner Seite, gibst mir das Gefühl dein Ein und Alles zu sein. Wir entdecken gemeinsam die Welt und bauen gleichzeitig zuhause unser Nest. Ich könnte mir niemand besseren an meiner Seite vorstellen. Danke, dass du mir jeden Tag zeigst, wie sehr du mich liebst. Ich bin wahnsinnig stolz, dich an meiner Seite zu haben und mit dir gemeinsame Erinnerungen zu schaffen. Wir sind ein Team und nur gemeinsam sind wir stark. Ich liebe dich. Von früher bis heute. Von jetzt bis immer.

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.

Ort, Datum Unterschrift