THE INFLUENCE OF GAZE-AUGMENTED REFLECTION ON STUDENTS' PROBLEM SOLVING

Axel Langner

DISSERTATION

The Influence of Gaze-Augmented Reflection on Students' Problem Solving

by
Axel Dinesch Langner

submitted in fulfillment of the requirements for the doctoral degree (Dr. rer. nat.) at the Faculty of Biology and Chemistry at the Justus-Liebig-University Giessen, Germany

Giessen, Germany, 2026

The results documented in this dissertation were obtained from April 2021 to October 2025 at the Institute of Chemistry Education at the Justus-Liebig-University Giessen under the supervision of Prof. Dr. Nicole Graulich.

The research project was funded by the German Research Foundation (DFG) within the project *How catchy are own eye-gaze replays? A Qualitative Study Exploring the Use of Students' Eye-Gaze Stimulated Retrospectives in Organic Chemistry – i-GAZE* (project number 453926661).

I would like to thank Prof. Dr. Nicole Graulich, the Institute of Chemistry Education, and the German Research Foundation (DFG). In particular, I am grateful to Prof. Dr. Nicole Graulich for her continuous support of my academic and professional development.

Det er ganske sandt, hvad Philosophien siger, at Livet maa forstaaes baglaends. Men derover glemmer man den anden Saetning, at det maa leves forlaends. Hvilken Saetning, jo meer den gjennemtaenkes, netop ender med, at Livet i Timeligheden aldrig ret bliver forstaaeligt, netop fordi jeg intet Øieblik kan faae fuldelig Ro til at indtage Stillingen: baglaends.

— Søren Kierkegaard (1843)

TABLE OF CONTENTS

ABSTRACT

Problem-solving activities support conceptual understanding and the development of competencies needed to resolve problems and make informed decisions. In organic chemistry, problem solving relies heavily on representations that encode both explicit and implicit information. However, students often struggle to use these representations in their problem solving.

Although various instructions have been shown to support problem solving with representations, they often benefit only specific subgroups of students. In contrast, reflection on one's own problem solving—enhanced by showing students their eye movements and providing guiding prompts in an eye-gaze-augmented retrospective—offers a differentiated and personalized approach. This dissertation investigated the influence of such gaze-augmented reflection on students' problem solving in two exploratory studies—with and without the retrospective.

This investigation integrates and extends the methodological and empirical foundations established in:

> Langner A. & Graulich N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804
>
> Langner A., Hain L., & Graulich N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, *102*(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830
>
> Langner A., Sahba M., Popova M., & Graulich N. (2025). An Integrated Approach to Characterizing Changes in Organic. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

The findings indicate that gaze-augmented reflection was associated with more goal-driven allocation of attention and that, despite highly individual trajectories shaped by initial problem-solving accuracy and student characteristics, accuracy converged across students. Overall, these findings highlight the potential of gaze-augmented reflection as a personalized approach to supporting students' problem solving with representations.

KURZFASSUNG

Problemlöseaufgaben fördern nicht nur das konzeptuelle Verständnis, sondern auch die Entwicklung zentraler Kompetenzen für das Lösen komplexer Probleme und das Treffen fundierter Entscheidungen. Insbesondere in der organischen Chemie spielen Repräsentationen, die sowohl explizite als auch implizite Informationen enthalten, eine zentrale Rolle. Studierende erleben jedoch häufig Herausforderungen dabei, diese Repräsentationen in ihren Problemlöseprozessen zu nutzen.

Obwohl verschiedene Instruktionsansätze das Potenzial zeigen, Studierende beim Problemlösen mit Repräsentationen zu unterstützen, profitieren davon häufig nur bestimmte Subgruppen. Demgegenüber könnte die Reflexion des eigenen Problemlöseprozesses, die durch die Einbindung eigener Blickbewegungen sowie gezielter Prompts im Rahmen einer blickbewegungsgestützten Retrospektive vertieft werden kann, eine differenzierte und individualisierte Lerngelegenheit darstellen.

Diese Dissertation untersucht den Einfluss einer solchen Reflexion auf die Problemlöseprozesse von Studierenden. Hierzu wurden zwei explorative Studien durchgeführt, eine mit und eine ohne Retrospektive. Die vorgestellte Untersuchung vereint und erweitert die methodischen und empirischen Grundlagen, die in den folgenden Arbeiten etabliert wurden:

> Langner A. & Graulich N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804
>
> Langner A., Hain L., & Graulich N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, *102*(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830
>
> Langner A., Sahba M., Popova M., & Graulich N. (2025). An Integrated Approach to Characterizing Changes in Organic. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

Die Ergebnisse zeigen, dass die Reflexion mit einer stärker zielgerichteten Aufmerksamkeitsverteilung verbunden ist und dass sich – trotz individueller Verläufe, die durch Antwortgenauigkeit und Studierendenmerkmale geprägt waren – die Antwortgenauigkeit zwischen den Studierenden annäherte. Insgesamt verdeutlichen diese Erkenntnisse das Potenzial der Reflexion in einer blickbewegungsgestützten Retrospektive als personalisierten Ansatz zur Unterstützung von Studierenden beim Problemlösen mit Repräsentationen.

LIST OF PUBLICATIONS

Publications in Peer-Reviewed Journals

Langner A., Sahba M., Popova M., & Graulich N. (2025). An Integrated Approach to Characterizing Changes in Organic. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

Langner A., Hain L., & Graulich N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, *102*(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830

Langner A. & Graulich N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804

Braun I., Langner A., & Graulich N. (2022). Let's draw molecules: Students' sequential drawing processes of resonance structures in organic chemistry. *Frontiers in Education*, 7, 1055280. https://doi.org/10.3389/feduc.2022.1055280

Langner A., Graulich N., & Nied M. (2022). Eye-Tracking as a Promising Tool in Pre-Service Teacher Education – A New Approach to Promote Skills for Digital Multimedia Design, *Journal of Chemical Education*, *99*(4), 1651–1659. https://doi.org/10.1021/acs.jchemed.1c01122

Eckhard J., Rodemer M., Langner A., Bernholt S., & Graulich N. (2021). Let's frame it differently – Analysis of instructors' mechanistic explanations, *Chemistry Education Research and Practice*, *23*(1), 78–99. https://doi.org/10.1039/D1RP00064K

Book Chapters

Bernholt S., Eckhard J., Rodemer M., Langner A., Asmussen G., & Graulich N. (2023). Designing tutorial videos to support students' learning of reaction mechanisms in organic chemistry. In Y. J. Dori, C. Nagi, & G. Szteinberg (Eds.), *Digital Learning and Teaching in Chemistry: An International and Inclusive Approach*, The Royal Society of Chemistry, 239–248. https://doi.org/10.1039/9781839167942-00234

Graulich N., Langner A., Vo K., & Yuriev E. (2021). Scaffolding metacognition and resource-activation during problem solving: A continuum perspective, In Tsaparlis G. (Ed.), *Problems and Problem Solving in Chemistry Education: Analysing Data, Looking for Patterns and Making Deductions*, The Royal Society of Chemistry, 38–67. https://doi.org/10.1039/9781839163586-00038

Conference Proceedings

Langner A. & Graulich N. (2024). Stimulating Reflection Processes with an Eye-Gaze-Augmented-Retrospective in Organic Chemistry, In Sardag, M., Kaya, G. & Sogut, M. (Eds.), *The 15th ESERA: Proceedings Book Series-1*, The 15th Conference of the European Science Education Research Association (ESERA), 252–258.

Langner A. & Graulich N. (2024). Mit Blick zurück einen Schritt vor: Eine Blickbewegungsgestützte Retrospektive. *Frühe naturwissenschaftliche Bildung*, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hamburg 2023, 186–189.

Braun I., Langner A. & Graulich N. (2023). Von Struktur zu Struktur: Untersuchung von Zeichenprozessen in der organischen Chemie, In van Vorst, H. (Ed.), *Lernen, Lehren und Forschen in einer digital geprägten Welt*, Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Aachen 2022, 486–489.

Langner A. & Graulich N. (2023). Lernförderlichkeit multimedialer Inhalte kriteriengeleitet abschätzen, In Wilke, T. & Rubner, I. (Eds.), *DiCE-Tagung 2023 - Digitalisation in Chemistry Education: Vol. Workshop*, DiCE-Tagung 2023 - Digitalisation in Chemistry Education, 1–3. https://doi.org/10.22032/dbt.59419

Oral Presentations

Langner A. & Graulich N. (2025). Neue Perspektiven. Ganzheitliche Charakterisierung von Blickänderungen. GDCP Jahrestagung 2025, Frankfurt, Germany.

Langner A. & Graulich N. (2025). Influence of an Eye-gaze-augmented Retrospetive on Students' Eye Movements and Success. ESERA 2025, Copenhagen, Denmark.

Langner A. & Graulich N. (2025). Eye-Tracking as a Feedback Tool in Education. Webinar Series of the Universitas Islam Negeri Sunan Kalijaga Yogyakarta (Indonesia), online.

Langner A. & Graulich N. (2024). Reflection on the Explanatory Power of Eye-Tracking Data. Group of Seminar of the Atkinson Group (University of North Texas), online.

Langner A. & Graulich N. (2024). Auf den zweiten Blick. Blickbewegungen zur Förderung von Reflexionen. GDCP Jahrestagung 2024, Bochum, Germany.

Langner A. & Graulich N. (2024). A second Glance. An Eye-Gaze-Augmented Retrospective to Stimulate Reflection on Problem-Solving in Organic Chemistry. Biennial Conference on Chemical Education 2024, Lexington, United States of America.

Langner A. (2024). Eye-Tracking und dessen Nutzung in der Forschung. Lehren Treffen Jahrgang 2019, Rauischholzhausen, Germany.

Langner A. (2024). How to communicate your research. First steps into science communication. Summer Symposium on Science Education 2024, online.

Langner A. (2024). Eye-Tracking in Science Education and novel ways to use it. Network in STEM Education (NiSE) University of Southern Denmark, online.

Langner A. & Graulich N. (2023). Eye-Tracking and Chemistry Education and a peek into my projects. Group of Seminar of the Popova Group (University of North Carolina Greensboro), online.

Langner A. (2023). Eye-Tracking in Chemistry Education and novel ways to use it. North Carolina Chemistry Education Research Division, online.

Langner A. & Graulich N. (2023). Mit Blick zurück einen Schritt vor. Eine blickbewegungsgestützte Retrospektive. GDCP-Jahrestagung 2023, Hamburg, Germany.

Langner A. & Graulich N. (2023). Stimulating Reflection Processes with an Eye-Gaze-Augmented Retrospective in Organic Chemistry. Center for Research in Science Education and Communication (University of Southern Denmark), Odense, Denmark.

Langner A. (2023). Kriteriengeleitetes Evaluieren der Lernförderlichkeit von Lehrvideos. Digitalisation in Chemistry Education, online.

Langner A. & Graulich N. (2023). Stimulating Reflection Processes with an Eye-Gaze-Augmented Retrospective in Organic Chemistry. ESERA 2023, Cappadocia.

Langner A. (2023). Multimediales Lernen und Lehren mithilfe von Eye-Tracking erfahren. Digitaltag 2023, Giessen, Germany.

Langner A. & Braun, I. (2023). Eye-Tracking in der chemischen Bildungsforschung. Gemeinsames Chemiedidaktik-Symposium der Universitäten Gießen & Münster, Münster, Germany.

Langner A. & Graulich N. (2022). How have I solved the problem? An eye-gaze augmented retrospective to foster students' comprehension of organic chemical representations. Biennial Conference on Chemical Education 2022, West Lafayette, United States of America.

Langner A. (2022) How Catchy are Own Eye-Gaze Replays? A Qualitative Study Exploring the Use of Students' Eye-Gaze Stimulated Retrospectives in Organic Chemistry. ESERA Summer School 2022, Utrecht, Netherlands.

Langner A., Nied M., & Graulich N. (2022). Multimediales Lernen und Lehren mithilfe von Eye-Tracking erfahren. 38. Fortbildungs- und Vortragstagung der GDCh-Fachgruppe Chemieunterricht, Osnabrück, Germany.

Langner A., Nied M., & Graulich N. (2022). Teaching digital multimedia design with eye-tracking. NARST: 2022 Annual International Conference, Vancouver, Canada.

Poster Presentations

Langner A. & Graulich N. (2024). Beim nächsten Mal… Halten sich Studierende an ihre eigenen Vorsätze bei zukünftiger Aufgabenbearbeitung. Jahrestagung der Fachgruppe Chemieunterricht, Regensburg, Germany.

Langner A. & Graulich N. (2023). Ein tiefer Einblick. Der Einsatz einer blickbewegungsgestützten Retrospektive zur Förderung von Reflexionsprozessen. GDCh Wissenschaftsforum Chemie, Leipzig, Germany.

Langner A. & Graulich N. (2023). Purposeful Reflection-on-Action: An Eye-Gaze-Augmented Retrospective to Foster Students' Comprehension of Organic Chemical Representations. Gordon Research Conference Chemistry Education Research and Practice 2023, Lewiston, United States of America.

Langner A. & Graulich N. (2022). Einflüsse Blickbewegungsgestützter Retrospektiven in der Organik. GDCP Jahrestagung 2022, Aachen, Germany.

Langner, A. & Graulich N. (2022). Wie betrachten Studierende Molekül-darstellungen beim Problemlösen? 38. Fortbildungs- und Vortragstagung der GDCh-Fachgruppe Chemieunterricht, Osnabrück, Germany.

Langner A. & Graulich N. (2022). How does problem- solving with organic chemistry molecules literally look like? Biennial Conference on Chemistry Education, West Lafayette, United States of America.

Langner A. (2022). How Catchy are Own Eye-Gaze Replays? A Qualitative Study Exploring the Use of Students' Eye-Gaze Stimulated Retrospectives in Organic Chemistry. ESERA Summer School 2022, Utrecht, Netherlands.

Langner A., Nied M., & Graulich N. (2022). Using eye-tracking to promote skills for multimedia design in pre-service teacher education. Royal Society of Chemistry (RSC). 2022 #RSCPoster Twitter Conference, online.

Langner A., Nied M., & Graulich N. (2021). Multimedia-Design im Lehramtsstudium Chemie – ein Seminarkonzept zur Evaluierung eigener Lehr-Lernmaterialien mithilfe von Eye-Tracking. Gesellschaft Deutscher Chemiker (GdCH). 37. Fortbildungs- und Vortragstagung der Fachgruppe Chemieunterricht, DiCE meets FGCU, online.

Langner A. & Graulich N. (2021). First Steps into Eye-Tracking in Chemical Education Research. Methods in Chemistry Education Research (MICER), online.

Langner A., Nied M., & Graulich N. (2021). Multimedia design in chemistry teacher training – A new teaching concept on evaluating one's own teaching material trough eye tracking. Variety in Chemistry Education / Physics Higher Education Conference (VICEPHEC), online.

Langner A. & Graulich N. (2021). Evaluating one's own teaching resources via eye-tracking in a student teacher training course. ESERA SIG Meeting Eye Tracking in Science Education Research, online.

Langner A., Eckhard J., & Graulich N. (2019). Ein Blick in den Spiegel. Studierende lernen von Studierenden beim Lösen von mechanistischen Problemen. Gesellschaft Deutscher Chemiker (GdCH). Wissenschaftsforum, Aachen, Germany.

HOW TO READ

Dear reader, although this dissertation is structured according to the conventions of academic research and grounded in established scholarly practices, its tone at times intentionally diverges from traditional academic language to make the work accessible beyond an academic audience, allowing readers without a research background to find orientation and meaning in the spirit of science communication. I hope you enjoy reading this dissertation, whether you are a researcher or simply interested in the topic.

THIS IS THE TITLE.

And this is the formal academic title commonly used in research.

This is regular text. Within this text, you will see black highlighted words that indicate subchapter titles. References are shown in grey, either within the text naming the authors (year) or in brackets (authors, year). You will also sometimes be directed to figures, tables, and equations (see Figure 0).

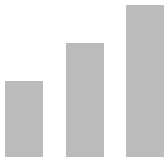


Figure 0 | This small text appears beneath a figure or an equation, or above a table, and provides a description or additional information.

SYNTHESIS: FOSTERING STUDENTS' PROBLEM SOLVING THROUGH AN EYE-GAZE–AUGMENTED RETROSPECTIVE

WHY?

(i.e., Introduction)

When individuals engage with a task for which no straightforward solution path is apparent, they engage in problem solving (Martinez, 1998). It is a complex cognitive activity that involves several interrelated steps, such as interpreting the given information and constructing an internal representation of the problem, identifying relevant strategies and prior knowledge, formulating hypotheses, verifying the solution, and learning from the process (cf. Garrett, 1987; Taconis et al., 2001; Weir, 1974).

Problem solving is particularly central in science education, where it supports the development of deep conceptual understanding and, through practice, fosters the acquisition of problem-solving competencies (Frey et al., 2022; Garrett, 1986). Ultimately, these competencies enable individuals to resolve problems and make informed decisions in science and related disciplines (Weir, 1974).

In organic chemistry, problem solving is closely linked to the use of representations. Organic chemists rely on symbolic-iconic representations that contain both explicitly visible features and implicitly present, non-visible features of chemical entities (see Figure 1) to reason about properties and associated chemical concepts, and to draw conclusions about chemical reactivity. To do so, they infer implicit information from the explicit, visually decodable representation of structural formulas (Goodwin, 2010; Hoffmann & Laszlo, 1991).

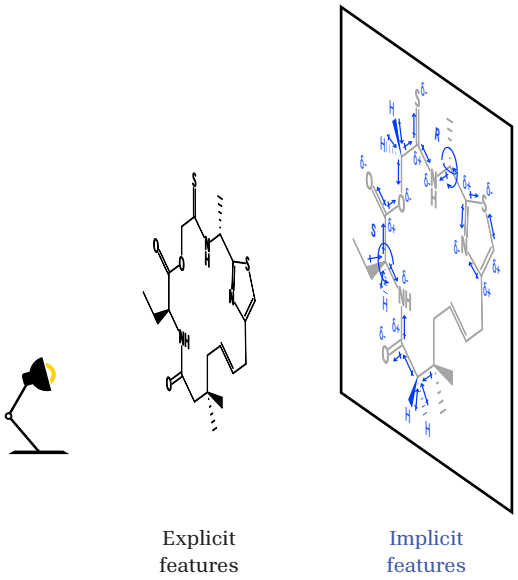


Figure 1 | Organic-chemistry representations depict explicitly visible features as well as implicitly present, but not explicitly visible, features (Hoffmann & Laszlo, 1991).

Consequently, students in organic chemistry need to develop expertise in solving problems using representations to advance toward professional practice. However, several studies have indicated that students often experience challenges when solving problems with representations. In these cases, they may struggle to derive

implicit properties from explicit structural features, as visually their focus is drawn to single, salient, or surface-level features or familiar patterns, and conceptually they may rely on intuitive judgments, memorization, misconceptions, or rule-based reasoning (cf. Anzovino & Bretz, 2015, 2016; Bhattacharyya & Bodner, 2005; Braun & Graulich, 2024a, 2024b; Braun et al., 2022; Braun et al., 2025; Christian & Talanquer, 2012; Cooper et al., 2013; Crowder & Raker, 2024; Cruz-Ramírez de Arellano & Towns, 2014; DeFever et al., 2015; Domin et al., 2008; Galloway et al., 2018; Graulich, 2015a; Graulich & Bhattacharyya, 2017; Graulich et al., 2019; McClary & Talanquer, 2011; Popova & Bretz, 2018a, 2018b; Rotich et al., 2024; Rushton et al., 2008; Sandi-Urena et al., 2019; Ward et al., 2025; Weinrich & Sevian, 2017).

Hence, instruction should support students in acquiring expertise in problem solving with representations. Studies across various science education disciplines have shown that instructional approaches such as example-based learning, process-oriented guided inquiry learning, contrasting cases, scaffolding, or productive failures can improve students' problem solving with representations (e.g., Kranz et al., 2025; Kranz et al., 2023; Stanford et al., 2018; van Gog & Rummel, 2010; Wernecke et al., 2018). However, across disciplines, uniform instructional approaches often benefit only a particular subgroup of students (cf. Hartmann et al., 2021; Kranz et al., 2025; Kranz et al., 2023; Lin et al., 2002; Reisslein et al., 2006; Sinha & Kapur, 2021; van Gog & Rummel, 2010; Wang et al., 2010).

In contrast, differentiated instructional approaches aim to provide personalized learning experiences that support students according to their individual needs, such as through reflection (cf. Dumont & Ready, 2023; Shemshack et al., 2021). Reflection can be described as an active and careful weighing of a belief or conviction by considering the reasons that underpin it and the further conclusions it leads to (Dewey, 1933). Across different perspectives (Boud et al., 1985; Boyd & Fales, 1983; Dewey, 1933; Killion & Todnem, 1991; Korthagen, 1985; Mezirow, 1991; Moon, 2004a; Sandars, 2009), it can be understood as a process with three stages: becoming aware or conscious of an experience, evaluating the experience, and ultimately revising existing beliefs or behaviors (cf. Atkins & Murphy, 1993; Koole et al., 2011; Peltier et al., 2005). Based on these three stages, several reflection-supporting models have been developed to guide reflective practice (e.g., Gibbs, 1988; Korthagen, 1985). One example is Korthagen's ALACT model (1985), developed for teacher education, which distinguishes five phases in reflection-based learning scenarios: action (A), looking back on the action (L), awareness of essential aspects (A), creating alternative methods of action (C), and trial (T).

The need for such models becomes particularly evident when reflection is conducted individually. When students reflect on their own, the reflection process may be limited, as they must assume both roles in an asymmetrical inner dialogue and may deceive themselves (Habermas, 1974). An educator who facilitates reflection can therefore encourage, accelerate, and deepen the reflective learning process (Boud et al., 1985) by using such models together with various methods, such as journals, role

plays, discussions, questions or prompts (cf. Bangert-Drowns et al., 2004; Guo, 2022; Nückles et al., 2020). In addition, technological tools can be used to record experiences for later reflection, such as videos, blogs, or digital portfolios (Hamel & Viau-Guay, 2019; Kori et al., 2014). Moreover, Fleck and Fitzpatrick (2010) argue that sensor devices can also be utilized to detect, capture, and represent information that would otherwise not be available to a student's memory, thereby offering an additional perspective for reflection.

In the context of supporting students' reflection, an eye tracker can be used as such a sensor device (Hansen et al., 2019; Szulewski et al., 2018). An eye tracker records where, for how long, and in which order an individual looks at a visual stimulus (Holmqvist et al., 2011). It outputs an eye-gaze replay that dynamically displays conscious and unconscious fixations (periods in which the eyes remain relatively still) and saccades (voluntary or reflexive rapid movements between fixations) (Duchowski, 2017; Holmqvist et al., 2011). This replay not only provides students with information about their visual processing of a complex representation, but also makes underlying cognitive actions more explicit (cf. Just & Carpenter, 1980; Lauwereyns & d'Ydewalle, 1996; Liu & Cui, 2025; Rayner, 1998; van Gog et al., 2005). In addition, it may reveal tacit knowledge (van Gog et al., 2005), activate cognitive resources (cf. Elby, 2000), and foster metacognition (Tsai et al., 2019).

Hence, a learning intervention—accompanied by an educator who uses guiding prompts and reflective questions while showing students their own eye-gaze replay in a retrospective (referred to as an eye-gaze-augmented retrospective)—may enable organic chemistry students to revisit their problem-solving approach in detail, provide them with additional individualized information, make them aware of crucial aspects of their unique process, and help them draw tailored consequences to enhance future problem solving with representations, whether in their visual or conceptual processing. Such an intervention offers individualized support. However, its use in this context has not yet been investigated. Therefore, two exploratory studies were conducted to examine the influence of such gaze-augmented reflection on students' problem solving: one in which students solved a problem-solving task, received an eye-gaze-augmented retrospective, and then solved a similar task again, and another in which students solved the two tasks consecutively without an eye-gaze-augmented retrospective in between.

WHAT?

(i.e., Research Purpose)

The purpose of this dissertation is to examine the influence of gaze-augmented reflection on students' problem solving in organic chemistry. This aim is operationalized by investigating the extent to which reflections stimulated by the eye-gaze-augmented retrospective (characterized by reflection complexity) shapes students' subsequent problem-solving performance (characterized by the change in accuracy and eye movements). This investigation builds on the methodological and empirical foundations established in Langner and Graulich (2024) (Chapter Students' Reflection), Langner, Hain, et al. (2025) (Chapter Defining Areas of Interest) and Langner, Sahba, et al. (2025) (Chapter Changes in Eye Movements).

Guided by theoretical perspectives on reflection, the following interrelated hypotheses frame the investigation to advance understanding of how students' engagement in an eye-gaze-augmented retrospective influences their problem solving:

1.) Students' initial problem-solving accuracy accounts for the complexity of their reflection.

2.) The complexity of students' reflection leads to changes in their performance (i.e., accuracy and eye movements) in subsequent problem-solving.

3.) The relationships between reflection complexity and changes in performance differ depending on the visual and conceptual affordances of organic chemistry tasks.

WITH?

(i.e., Research Design)

In study I, students received the eye-gaze-augmented retrospective between each pair of tasks (see Figure 2).

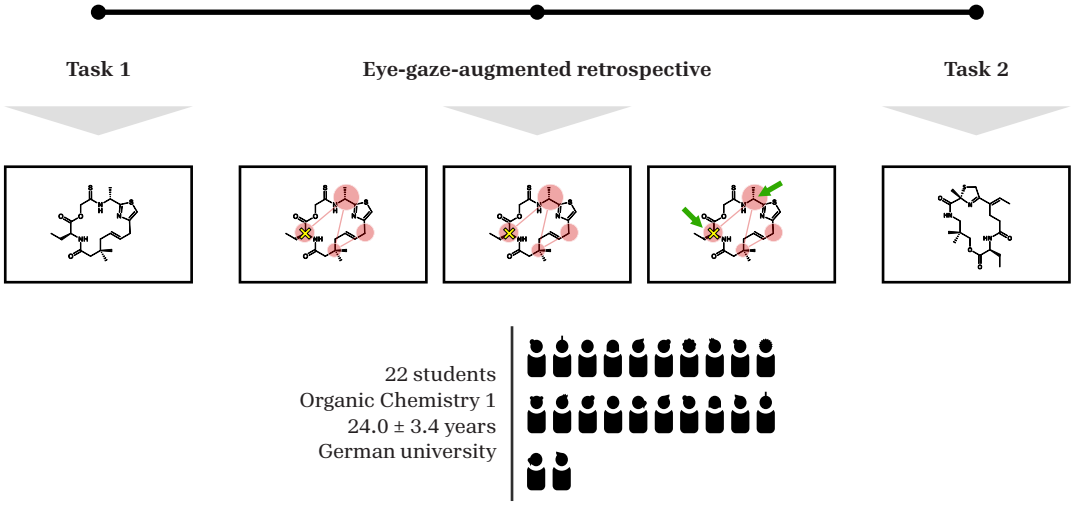


Figure 2 | Study I. Data were collected in single eye-tracking and interview sessions in 2021 at a German university. 22 students were recruited voluntarily from an Organic Chemistry 1 lecture. The lecture followed a traditional curriculum structured around functional groups, mechanism types, and structure–property relationships, with Buddrus and Schmidt (2015) serving as the primary textbook. The Students did not receive any form of compensation for their participation in the study. Sixteen students identified as female and six as male. The average age of the students was 24.0 ± 3.4 years. All students were native German speakers with normal or corrected-to-normal vision.

In contrast, in study II, students solved the tasks in direct succession without an intervention in between (see Figure 3). Relative to Study I, this design offers insight into students' problem solving when no eye-gaze-augmented retrospective is provided.

In both studies, the problem-solving task contexts of R/S configuration of asymmetric carbon atoms, Brønsted acid–base concept, and electrophilicity and nucleophilicity were chosen as they cover a wide range of visual affordances (i.e., task-relevant visually represented features) and conceptual affordances (i.e., rules or principles underlying the chemical concepts queried by the task) relevant to beginner-level organic chemistry. For each context, two similar tasks with the same prompt but different representations were designed (see Figure 4).

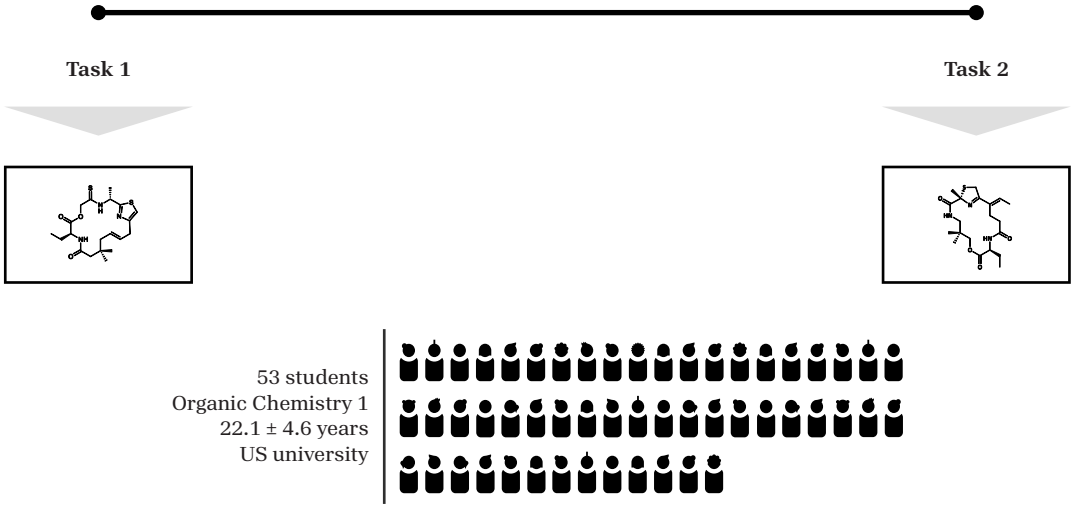


Figure 3 | Study II. Data were collected in single eye-tracking sessions in 2023 at an US university. 53 students were recruited voluntarily from two Organic Chemistry 1 lectures. The lectures followed a traditional curriculum and were structured by functional groups, mechanism types, and structure-property-relationships, with Klein (2020) serving as the primary textbook. As compensation, students received extra credit toward their final exam grades. They could earn these bonus points either by participating in this study or by completing an alternative assignment. Forty-four students identified as female and nine students as male. The average age of the students was 22.1 ± 4.6 years. The students were native English speakers who had normal or corrected-to-normal vision.

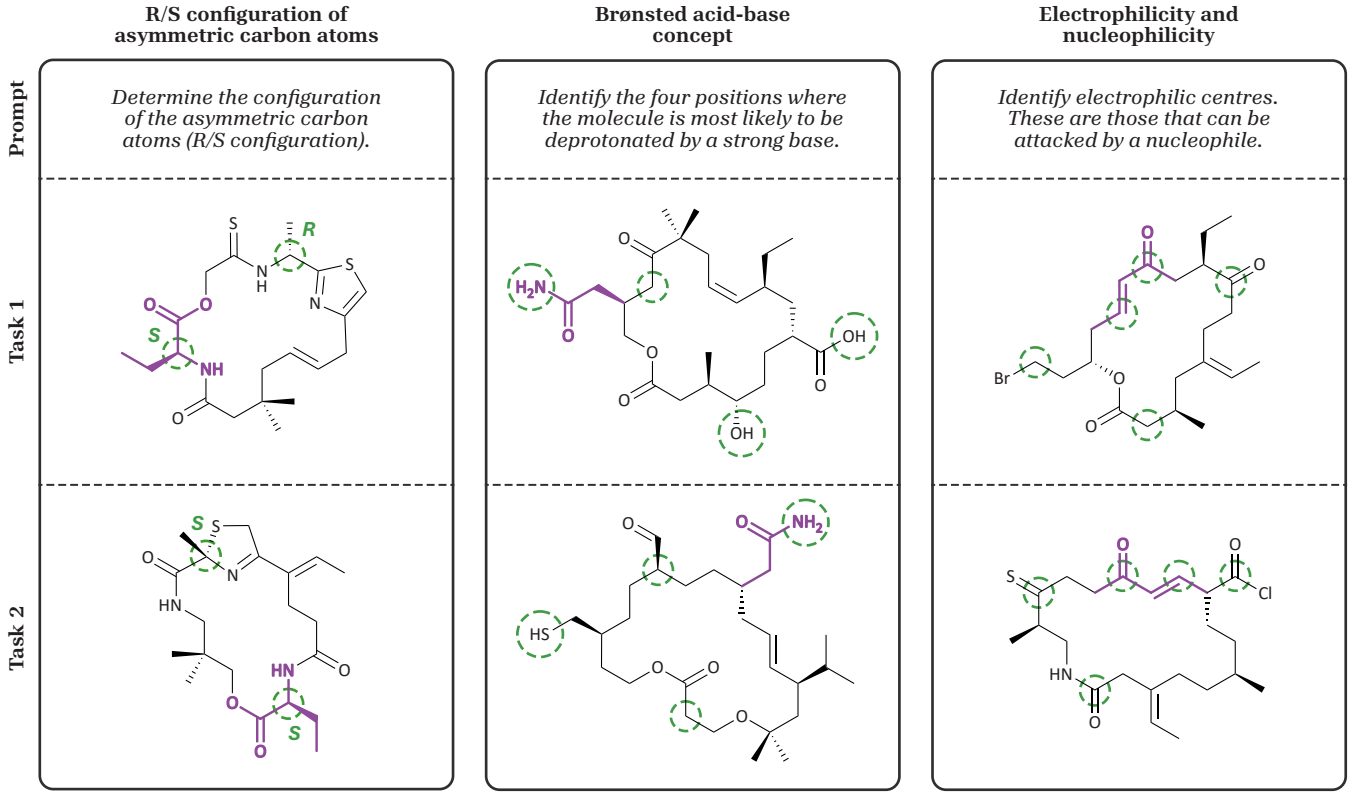


Figure 4 | Task design used in both studies. The task solution is shown in green. The molecular structures were large cyclic organic molecules resembling macrolides (cf. Arsic et al., 2018). Identical structural sites were incorporated into the representations of each task pair (highlighted in purple), ensuring that the representations were identical at a local level and enabling a comparison of students' eye movements at specific key structural sites, while remaining distinct at a global level, thereby minimizing the risk of practice effects.

The eye-gaze-augmented retrospective designed for study I (see Figure 2) followed the reflection phases of Korthagen's (1985) ALACT model and incorporated guidelines, considerations, and recommendations from Davis (2003), Ericsson and Simon (1993), Fleck and Fitzpatrick (2010), Hyrskykari et al. (2008), King (1994), Kori et al. (2014), Korthagen and Kessels (1999), Moon (2004a, 2004b), and Sandars (2009) (see Figure 5). For a detailed description of the design of the eye-gaze-augmented retrospective, see Chapter Students' Reflection.

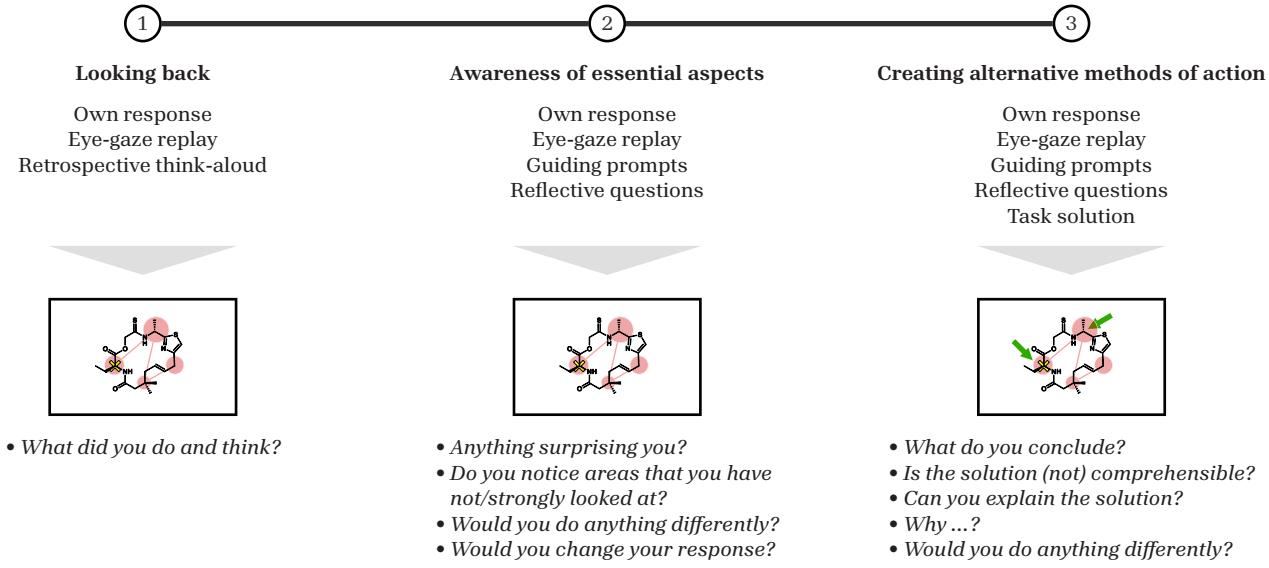


Figure 5 | The eye-gaze-augmented retrospective, embedded between two similar tasks and implemented as a semi-structured interview. The prompts and questions shown here are shortened English versions; the full German versions are provided in Appendix A.

After solving a problem (action), students watched an eye-gaze replay of their own problem-solving process, superimposed with their response, and were instructed to describe this process (looking back). Beforehand, they received an explanation of fixations and saccades as displayed in the replay. Students were also informed that they could pause, rewind, fast-forward, or rewatch the replay as often as they wished. Prompts and questions supported students in evaluating their problem-solving process (awareness of essential aspects). Next, the task solution was superimposed in the replay, and further prompts and questions guided students to consider alternative approaches and draw consequences for future problem solving (creating alternative methods of action). All guiding prompts and reflective questions aim to direct students' attention to key events, activate additional cognitive resources, and make tacit knowledge explicit and available for reflection (the original prompts and questions in German are provided in Appendix A). Finally, students solved a similar problem again (trial).

HOW?

(i.e., Data Collection and Analysis)

For the data collection in both studies, ethical standards were followed in accordance with the 2013 Declaration of Helsinki and its later amendments or comparable ethical guidelines for educational research. At the German university, Institutional Review Board (IRB) or comparable ethics approval is not required for chemistry education research. Nevertheless, a rigorous internal review protocol was implemented to ensure ethical integrity for study I. For study II, all recruitment, data collection, and data analysis procedures were reviewed and approved by the Institutional Review Board (IRB-FY24-113). Participation in both studies was voluntary, and all participants provided written informed consent after receiving information about the study's purpose, their rights, and data handling procedures. No identifiable personal information is reported in this dissertation. All data collection procedures were conducted in the students' respective native languages.

In both studies, students first completed a short demographic questionnaire (for the original questionnaires, see Appendix B).

While solving the tasks (see Figure 4), eye-tracking data of students were collected. First, the task prompt was presented on a monitor. Once students had read and understood it, they continued by pressing a key. A fixation cross then appeared for two seconds before the molecular representation was automatically displayed. After completing the task, the students ended it by pressing a key again (see Figure 6). The original stimuli presented to students are provided in Appendix C.

For the accuracy data collection, students recorded their responses on a sheet of paper depicting the respective molecule after solving the problem mentally in front of the eye-tracker.

The interview data of the eye-gaze-augmented retrospective in study I (see Figure 5) was audio taped and transcribed verbatim.

Additionally, supplementary data on organic chemistry knowledge with self-rating items at the beginning of the main data collection and perceived confidence, task difficulty, mental load, and cognitive load were collected using items from Brandriet & Bretz (2014), Bratfisch et al. (1972), Hart & Staveland (1988), Hart (2006), Klepsch et al. (2017), and Leppink et al. (2013) after each task were collected. These measures were originally included for potential exploratory purposes but were not used in the final analysis and do not influence the reported results.

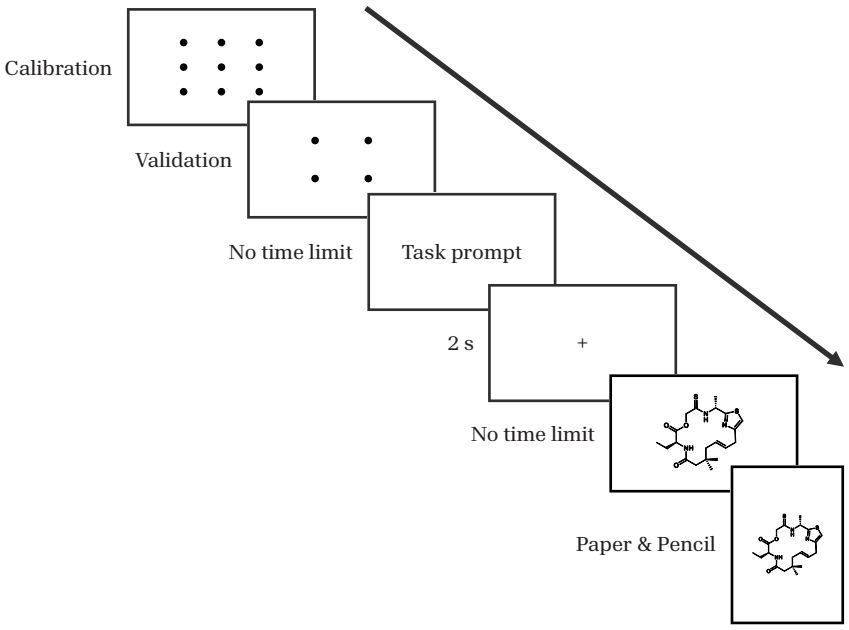


Figure 6 | Eye-tracking data collection procedure. Eye movements were recorded using a Tobii Pro X3-120 screen-based eye tracker (120 Hz sampling rate) in combination with a 24-inch presentation monitor, a Tobii Pro External Processing Unit, and the accompanying software Tobii Pro Lab in a live-viewer setup. Participants were seated 60–65 cm from the monitor and were instructed to minimize head and body movements to reduce data loss. A nine-point calibration followed by a four-point validation was performed prior to each recording. The molecule was scaled so that the distances between atoms corresponded to an eye movement of approximately 2°, exceeding both the system's accuracy (0.7°) and the 1–1.5° recommended by Holmqvist et al. (2011), allowing fixations to be reliably assigned to viewed atoms.

For the subsequent data analysis, only eye-tracking data, accuracy data, and interview data associated with the key structural sites that were identical across each task pair were considered (see Figure 4), as they represent positions at which students' eye movements can be directly compared across the two tasks. Additionally, observations of task pairs in which eye-tracking data loss exceeded 20% were excluded from analysis to ensure data quality. After exclusion, 57 observations for study I and 130 for study II across all task pairs remained.

The complexity of reflections was characterized by the presence or absence of reflection elements derived from models and theories of reflection (Atkins & Murphy, 1993; Boyd & Fales, 1983; Fleck & Fitzpatrick, 2010; Hatton & Smith, 1995; Kember, 1999; Kember et al., 2000; Koole et al., 2011; Korthagen, 1985; Korthagen & Kessels, 1999; Mezirow, 1981, 1990, 1991; Moon, 2004a; Nguyen et al., 2014; Peltier et al., 2005) (see Table 1). These elements in students' reflections were coded dichotomously using deductive qualitative content analysis. Based on the co-occurrence of the reflection elements, the reflection processes were categorized as either lower or higher in complexity (for a fine-grained discussion of reflection complexity, see Chapter Students' Reflection). While reflections of lower complexity primarily involved describing experiences and, in some cases, identifying difficulties or challenges that students encountered or still faced, reflections of higher complexity additionally included evaluating different approaches or drawing consequences (see Table 2).

Table 1 | Description and examples of the codes, representing reflection elements, used to characterize the structural complexity of the reflection processes.

Code	Code description	Example
Describing the experience (E)	The participant describes or recalls the problem-solving process (e.g., related to chemical concept knowledge, visual behavior, or general thoughts).	"I focused on where the hydrogen atoms are. [...] And at first I had to think about it a bit."
Identifying an issue (I)	The participant verbalizes difficulties, challenges, or uncertainties they had or still have (e.g., related to chemical concept knowledge, visual behavior, or general behavior).	"But then I felt kind of uncertain, because I thought, I don't think I've ever consciously seen something like NH^- or anything like that before."
Evaluating an approach (A)	The participant reviews their assumptions and behavior, or tests alternative assumptions and behavior (e.g., related to chemical concept knowledge, visual behavior, or general behavior).	"Maybe it can kind of flip over somehow, and then move up to that. Like, the double bond could flip up [...]"
Drawing a consequence (C)	The participant verbalizes intentions or changes regarding their actions and thoughts with reference to future problem-solving situations (e.g., related to chemical concept knowledge, visual behavior, general behavior, or general thoughts).	"Next time, I would take more time and look at the whole molecule more carefully."

Table 2 | Categorization of reflection processes as lower or higher in complexity based on the presence or absence of coded reflection elements.

Reflexion complexity	Description
Reflexions of higher complexity	The reflection process involves describing experiences (E) and issues (I), evaluating one's own or an alternative approach (A), and verbalizing intentions or changes with reference to future problem solving (C).
	The reflection process involves describing experiences (E) and issues (I), and additionally verbalizes intentions or changes with reference to future problem solving (C).
	The reflection process involves describing experiences (E) and issues (I), and additionally evaluates problem solving by reviewing or testing an alternative approach (A).
Reflexions of lower complexity	The reflection process involves describing experiences (E) and additionally expresses issues encountered during problem solving or that still persist (I).
	The reflection process involves describing experiences (E).

The accuracy of students' responses were scored. For the R/S configuration of asymmetric carbon atoms task set, students received a score of 1 if they accurately identified the chiral center and accurately determined the configuration, a score of 0.5 if they accurately identified the chiral center but inaccurately determined the configuration, and a score of 0 if they did not identify the chiral center. For the Brønsted acid-base concept task set, students received a score of 1 if they accurately identified the nitrogen of the carboxamide and a score of 0 if they did not. For the

electrophilicity and nucleophilicity task set, students received a score of 1 if they accurately identified the β-carbon of the α,β-unsaturated carbonyl site and a score of 0 if they did not. Students' change in accuracy was calculated as the difference in accuracy scores between the two tasks in each task pair.

To characterize the overall change in eye movements, multiple eye-tracking metrics were included in the analysis. Using several metrics allows the capture of complementary characteristics of eye movements, thereby offering a more comprehensive picture of students' information-processing behavior (cf. Holmqvist et al., 2011; Skaramagkas et al., 2021). Prior research has demonstrated the value of several metrics for characterizing various aspects of chemistry problem solving such as students' strategies, their use of representations, and the cognitive demands of tasks (Baluyut & Holme, 2019; Braun et al., 2022; Cullipher & Sevian, 2015; Hinze et al., 2013; Karch et al., 2019; Locatelli et al., 2024; Rodemer et al., 2023; Tang et al., 2014; Tang & Pienta, 2012; Tóthová et al., 2021; Williamson et al., 2013; Ye et al., 2024). Hence, multiple metrics were aggregated by calculating the average absolute bounded normalized change across several eye-tracking metrics Δ_{Gaze} between each task pair according to Equation 1 (see Chapter Changes in Eye Movements).

$$\Delta_{\text{Gaze}} = \frac{1}{\text{number of metrics}} \sum_{\text{metric}=1}^{\text{number of metrics}} \left(\frac{1}{\text{number of subdimensions}_{\text{metric}}} \sum_{\text{subdimension}=1}^{\text{number of subdimensions}_{\text{metric}}} \left| \frac{\text{value}^{\text{task2}}_{\text{metric,subdimension}} - \text{value}^{\text{task1}}_{\text{metric,subdimension}}}{\text{value}^{\text{task2}}_{\text{metric,subdimension}} + \text{value}^{\text{task1}}_{\text{metric,subdimension}}} \right| \cdot \mathbf{1}\left[\text{value}^{\text{task2}}_{\text{metric,subdimension}} + \text{value}^{\text{task1}}_{\text{metric,subdimension}} > 0\right] \right)$$

Equation 1 | The average absolute bounded normalized change across eye-tracking metrics Δ_{Gaze}, a statistical composite calculation to determine the change in students' eye movements through an aggregated measure.

For this, Areas of Interest (AOIs) for the molecules in each task were defined using the limited-radius Voronoi tessellation method (see Chapter Defining Areas of Interest) (see Figure 7).

Next, the metrics fixation duration, fixation count, AOI hits, pattern frequency, transitions, transition-fixation ratio, transition entropy, and fixation disparity were calculated, reflecting spatial, temporal, and psychophysiological information in eye movements (cf. Holmqvist et al., 2011; Just & Carpenter, 1980; Koh et al., 2024). For a descriptions of these metrics, see Chapter Changes in Eye Movements. Redundant metrics were excluded through correlation and multicollinearity analysis of related metrics. For each task pair (see Figure 4), data from the first task were used for these analysis as it serves as a point of origin. The resulting set of included metrics for each task pair is shown in Table 3.

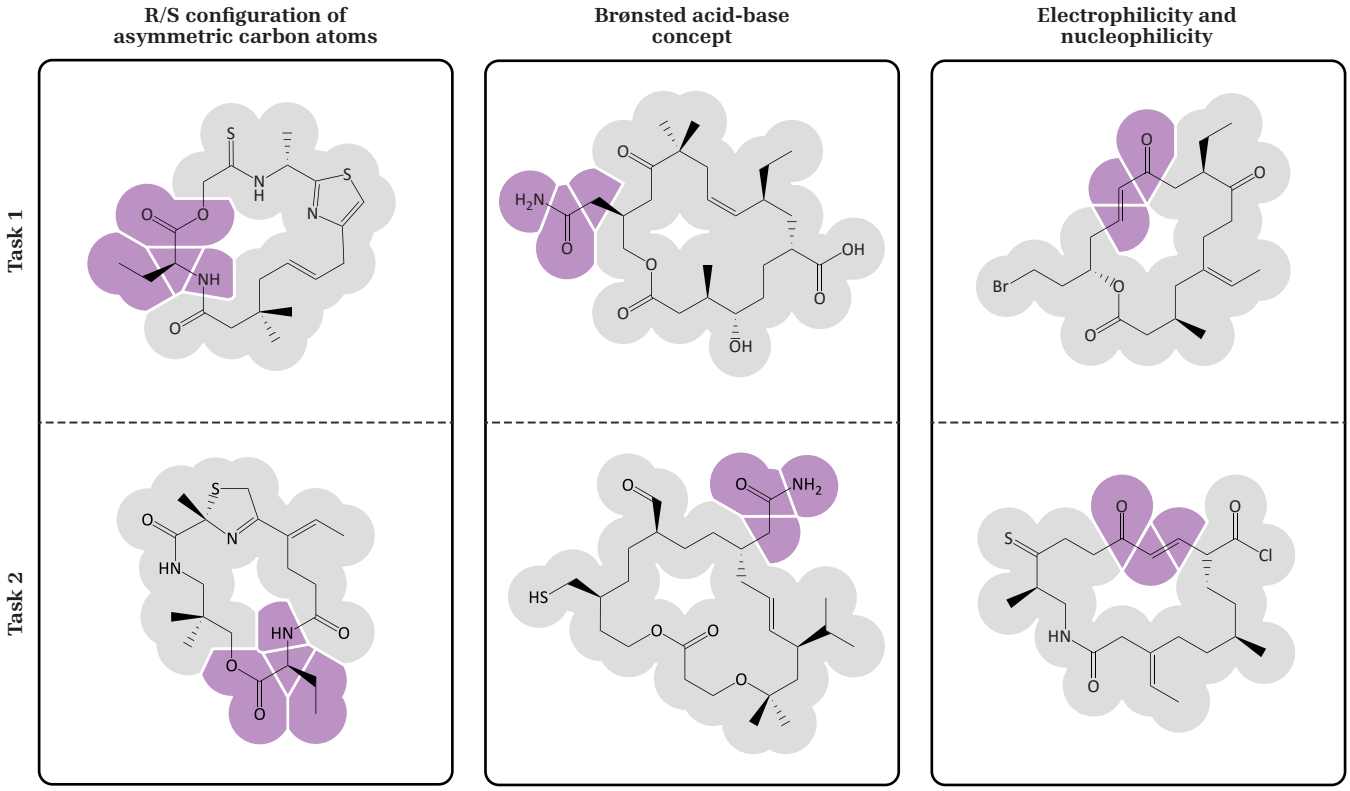


Figure 7 | AOIs defined using the limited-radius Voronoi tessellation method, with the radius set to the bond length of the molecules. The AOIs within the key structural sites were merged to correspond to their structural features (highlighted in purple), whereas the remaining portions of the molecule were combined into a single, larger AOI (highlighted in grey).

Table 3 | Included metrics in the calculation of Δ_{Gaze} for each pair of tasks.

R/S configuration of asymmetric carbon atoms	Brønsted acid-base concept	Electrophilicity and nucleophilicity
Absolute total fixation duration, relative total fixation duration, absolute pattern frequency[a], relative pattern frequency[a], transition entropy, transition-fixation ratio, fixation disparity	Absolute total fixation duration, relative total fixation duration, relative transitions frequency, absolute pattern frequency[b], transition entropy, transition-fixation ratio, fixation disparity	Absolute total fixation duration, relative total fixation duration, relative transitions frequency, absolute pattern frequency[b], transition entropy, transition-fixation ratio, fixation disparity

[a]three-AOI sequence reflecting the specific sequential events of switching between the residues of the carbon center, and switching between carbon center and residues
[b]three-AOI sequence reflecting the specific sequential events of switching between all AOIs, and switching between two specific AOIs (e.g., the carbonyl and the α-carbon)

SO?

(i.e., Results and Discussion)

To examine whether students' initial problem-solving accuracy accounts for the complexity of their reflection, the analysis focuses on how reflection complexity develops across the phases of the retrospective in relation to students' accuracy. The following examples of Mary's, Kate's and Susan's reflection illustrate this relationship (see Figures 8, 9, and 10).

In Mary's reflection of lower complexity (see Figure 8), she described her problem-solving approach and identified an issue she encountered. She was unsure whether the nitrogen of the carboxamide would donate a hydrogen. However, Mary's reflection process concluded once her response was confirmed as accurate, despite her stated uncertainty. Her reflection illustrates a pattern seen among students who recognize issues but do not further examine the underlying reasons of it or extend their reflection beyond acknowledging the issue (for another example, see Gwen's reflection in Chapter Students' Reflection). Since the experienced issue did not affect their accurate response, those students may not perceive a need to explore the cause of it in their reflection—even though further reflection might benefit them. Confirmation of their response could have created the impression that external correctness equates to internal understanding, a phenomenon consistent with the Illusion of Explanatory Depth (Rozenblit & Keil, 2002). Moreover, having one's response validated might retrospectively transform uncertainty into a sense of certainty, producing cognitive closure (Kruglanski & Webster, 1996) or even hindsight effects, in which students believe they knew it all along (Fischhoff, 1975). In these cases, the

Mary's reflection of lower complexity

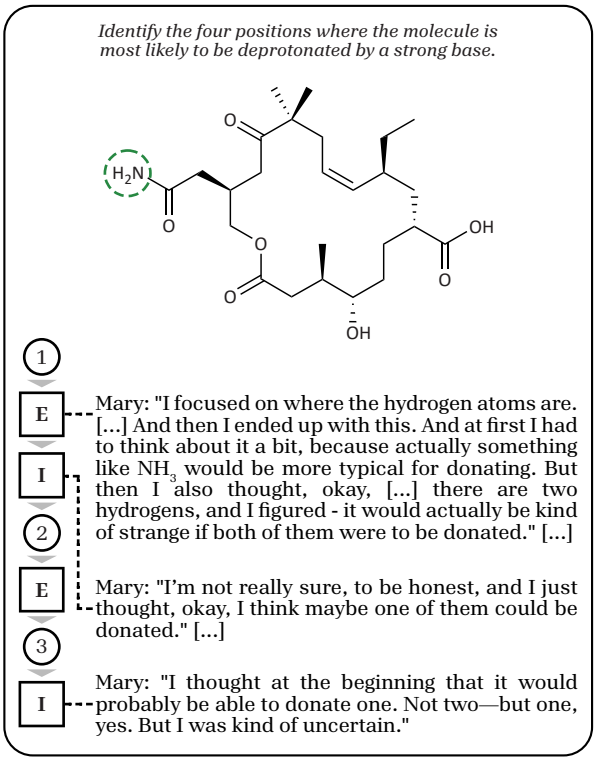


Figure 8 | Excerpt from Mary's reflection of lower complexity at the nitrogen of the carboxamide of the task about Brønsted acid-base concept. The numbers correspond to the phases of the retrospective (see Figure 5), and the letters correspond to the reflection elements (see Table 1).

retrospective did not stimulate a reflection of higher complexity, even though uncertainty (Dewey, 1933; Schön, 1983)—along with related experiences such as doubt, confusion, concerns about not knowing, and lacking sufficient information to explain something (for a review, see Marshall et al., 2021)—can serve as a stimulus for further reflection, as illustrated in Kate's reflection of higher complexity (see Figure 9).

In the beginning, Kate also expressed uncertainty. After seeing the task solution, she generated a hypothesis to explain why the solution was correct. Building on this assumption, she explored the underlying reason for her inaccurate response and ultimately concluded that, in future problem-solving situations, she should take more time and examine the representation more carefully.

Although both Mary and Kate identified an issue, their reflections differ in complexity. This difference emerges in the third phase of the retrospective, when students verify whether their response was accurate (see Figure 5).

Kate's reflection of higher complexity

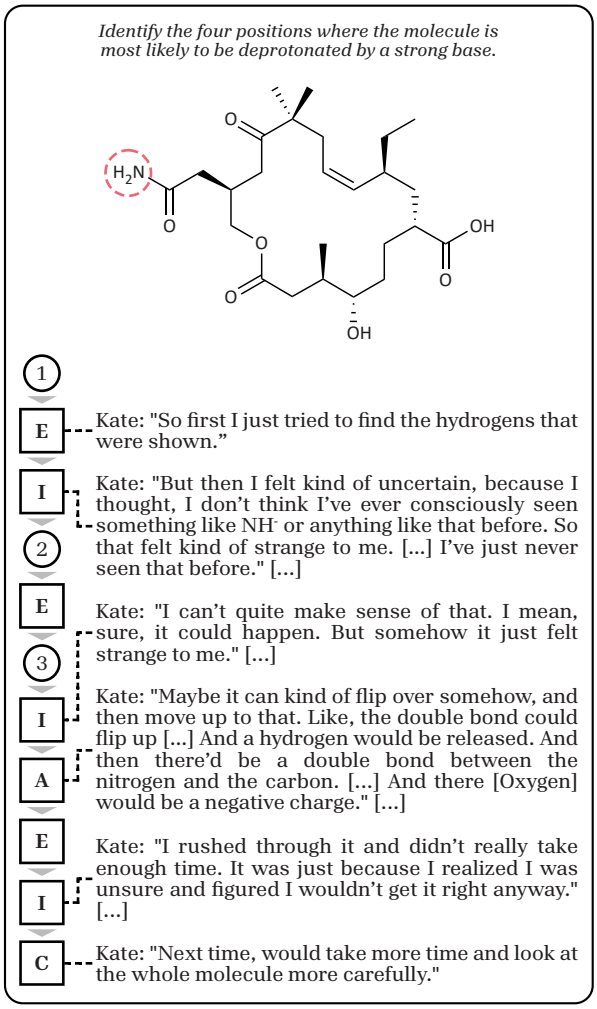


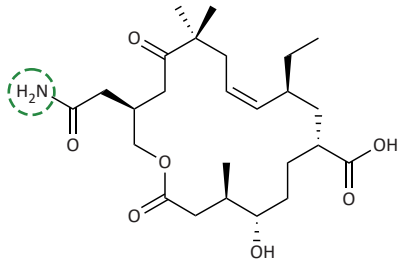
Figure 9 | Excerpt from Kate's reflection of higher complexity at the nitrogen of the carboxamide of the task about Brønsted acid-base concept.

However, as illustrated by Susan's reflection (see Figure 10), students did also exhibit higher reflection complexity regardless of the accuracy feedback in the retrospective. Susan drew the consequence that she needed to review the Brønsted acid–base concept—even before the task solution was revealed to her (for another example, see Felicia in Chapter Students' Reflection).

This illustrates that while uncertainty and related experiences can trigger reflections of higher complexity, uncertainty alone is not always sufficient. In such cases, additional experiences of inner discomfort, surprise, confusion, an unwanted outcome, difficulty in understanding, or a desire to make sense of the content—known stimuli for reflection (Boyd & Fales, 1983; Mezirow, 1991; Moon, 2004a; Schön, 1983)—elicited by receiving feedback that a structural site was determined inaccurately may have prompted the emergence of reflections of higher complexity.

Susan's reflection of higher complexity

Identify the four positions where the molecule is most likely to be deprotonated by a strong base.



1

E — Susan: "Okay, Brønsted. I just went by the question, what can actually be deprotonated?" [...]

I — Susan: "I wasn't sure about that. I think, actually, it's more likely to be protonated than deprotonated." [...]

E — Susan: "But at first glance, okay, there's definitely an H that could just come off." [...]

2

E — Susan: "But now, looking back afterwards — you'd actually rather have NH_3^+. If anything happens there."

A — Susan: "So yeah, that was just intuition in the first moment."

E — Susan: "But then, thinking about it afterwards... Well, the concepts of Brønsted bases, I just don't really have them right now — apart from proton donor and acceptor." [...]

I — Susan: "So I'd basically go back and review the whole Brønsted acid/base concept again."

C — Susan: "I don't think I would check that option again now." [...]

A

3

C — Susan: "I really need to look at the Brønsted concept again, definitely."

Figure 10 | Excerpt of Susan's reflection of higher complexity at the nitrogen of the carboxamide of the task about Brønsted acid-base concept.

How the students' initial accuracy can shape the complexity of students' reflection is shown in Figure 11.

In the first two phases of the retrospective, the distribution of lower and higher reflection complexity remained similar, with lower complexity predominating, regardless of whether students had determined the structural site accurately or not. However, in the third phase, a notable shift emerged. Although the proportion of higher reflection complexity increased, the increase was substantially greater for reflections on inaccurately determined structural sites (from 22% to 89%) than for reflections on accurately determined sites (from 33% to 63%). When considering a broader set of students' reflections, this pattern becomes even more pronounced (see Chapter From Sight to Insight).

Overall, these findings suggest that reflections of higher complexity are stimulated when students not only identify an issue but also faced with an inaccurate response (for example, see Kate's reflection in Figure 9). Yet identifying an issue alone may or may not lead to more complex reflection—in some cases it does (for example, see Susan's reflection in Figure 10), whereas in others it does not (for example, see Mary's reflection in Figure 8).

To examine whether the complexity of students' reflection leads to changes in their performance (i.e., accuracy and eye movements) in subsequent problem-solving, students' changes in score and their eye movements across the two tasks of each task pair were compared as a function of their reflection complexity. The distribution of score changes in relation to reflection complexity in study I revealed three main observations (see Figure 12, study I).

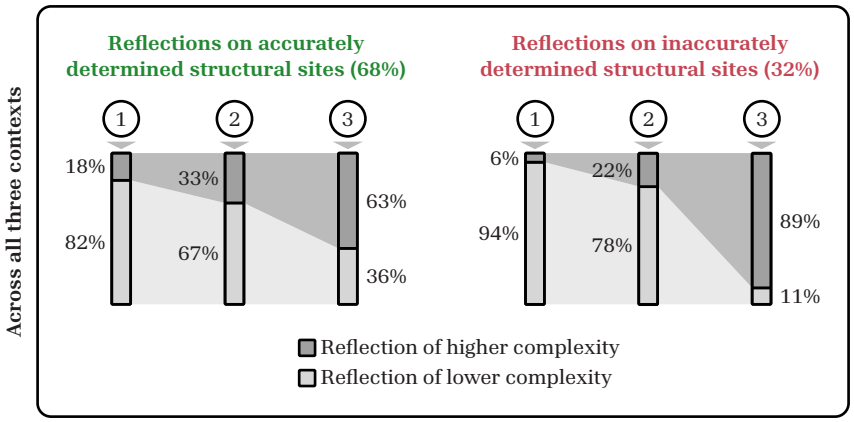


Figure 11 | Occurrence of reflection complexities across the three phases of the retrospective for accurately (left) and inaccurately (right) determined structural sites across all three contexts. The numbers indicate the percentage distribution of reflection complexities within each phase.

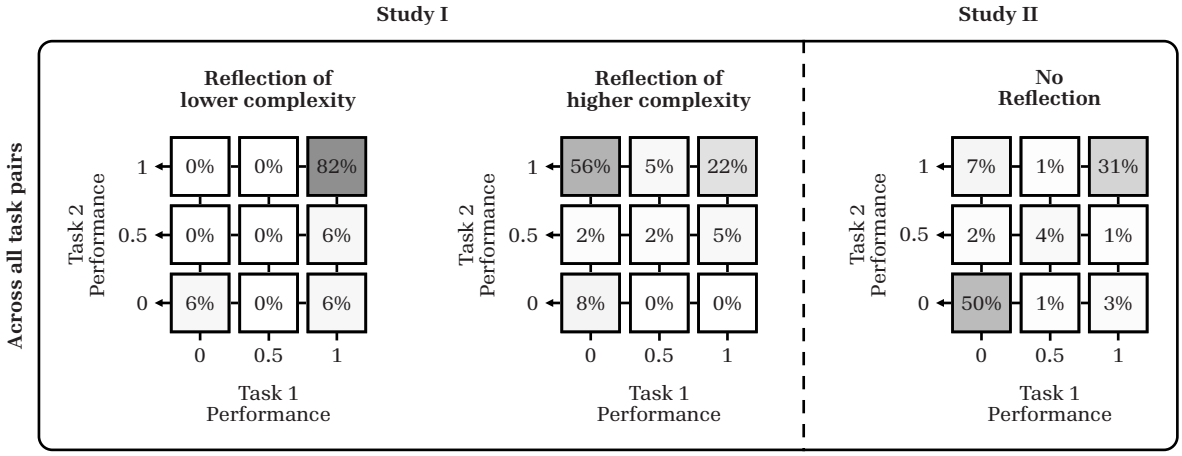


Figure 12 | Change of students' score across all task pairs as a function of reflection complexity in study I (left). For comparison, the change of students' score of study II are shown (right). Note: The score of 0.5 was only given in R/S configuration of asymmetric carbon atoms task set, leading to that this score is underrepresented.

First, students accurately determined the structural site in the initial task (score of 1), reflected with lower complexity, and again responded accurately in the subsequent task (score of 1) (see Figure 12, reflection of lower complexity; 82%). Many of these students—such as James (see Figure 13)—were pleased with their performance and appeared to see no need to extend their reflection (for another example, see Steven in Chapter Students' Reflection). They effectively judged that further reflection was unnecessary, likely relying on past experiences, response time, and fluency (as reviewed by Bjork et al., 2013). In such cases, task demands aligned well with students' knowledge. Their brief reflections demonstrated efficient metacognitive monitoring and control rather than superficiality (Son & Schwartz, 2002). In contrast, other students recognized issues but still did not explore them further—for example, Mary (see Figure 8) (for another example, see Gwen in Chapter Students' Reflection). Yet their issues did not impair their accuracy at all. This suggests that when an issue

does not threaten students' problem-solving accuracy, they may not engage in more complex reflection, showing a rather task-oriented behavior in which performance is judged against norm-referenced standards (Meece, 1994; Nicholls & Miller, 1984).

Second, students accurately determined the structural site in the initial task (score of 1), reflected with higher complexity, and again responded accurately in the subsequent task (score of 1) (see Figure 12, reflection of higher complexity; 22%). This suggests that the eye-gaze-augmented retrospective has the potential to support students—such as Susan (see Figure 10)—in engaging in more complex reflection that involves both metacognitive monitoring and regulation, even when they are already able to solve the problems accurately and consistently. Such students showed a learning-oriented behavior, valuing the development of their competence based on self-referenced standards (Conroy et al., 2009; Meece, 1994).

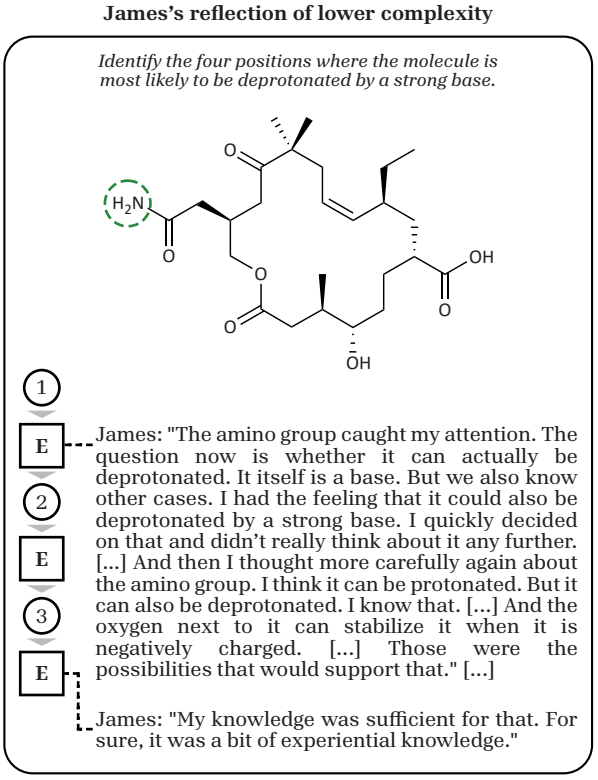


Figure 13 | Excerpt from James's reflection of lower complexity at the nitrogen of the carboxamide of the task about Brønsted acid-base concept.

Third, some students initially determined the structural site inaccurately (score of 0 or 0.5), reflected with higher complexity, and then accurately determined the subsequent task (score of 1) (see Figure 12, reflections of higher complexity; 61%). In contrast to cases like Susan (see Figure 10), for some students reflection of higher complexity was prompted by the accuracy feedback. For example, although Kate (see Figure 9) had already identified an issue in the beginning of her reflection, she extended her reflection only after the task solution revealed that her response was inaccurate—showing a task-oriented behavior similar to Mary (see Figure 8). In other cases, students were not initially aware of any issue and became aware of it only once the task solution revealed that their response was inaccurate. For example, after seeing the solution, Yelena immediately recognized that her exclusive focus on the carbon skeleton had caused her to not consider the carboxamide (see Figure 14) (for other examples, see Peter and Anna in Chapter Students' Reflection). Such moments illustrate an illusion of knowing, in which students believe they understand the content until feedback reveals otherwise (Glenberg et al., 1982). Besides the uncertainty, doubt, confusion, concerns about not knowing, or the recognition of insufficient explanations triggered by negative feedback that can stimulate more complex reflection (Dewey, 1933; Marshall et al., 2021; Schön, 1983), such

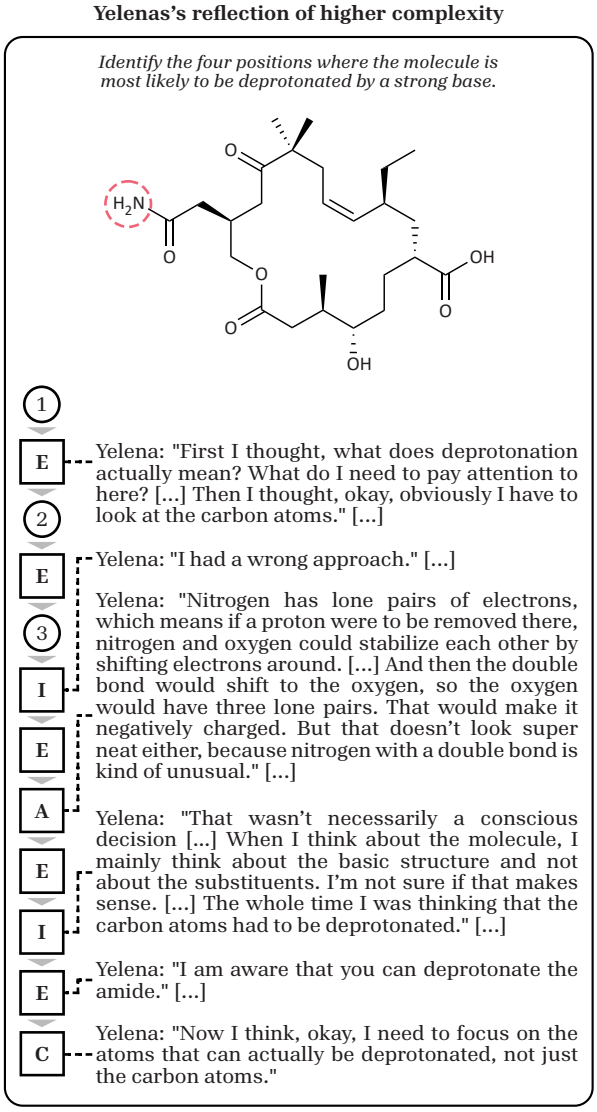


Figure 14 | Excerpt of Yelena's reflection of higher complexity at the nitrogen of the carboxamide of the task about Brønsted acid-base concept.

dissatisfaction with one's prior conceptions may also prompt willingness to invest effort (van Gog et al., 2024) and conceptual change (cf. Posner et al., 1982), as illustrated by Yelena's case. Even though those students were initially inaccurate, the eye-gaze-augmented retrospective guided them toward improved accuracy in the subsequent problem-solving task.

Overall, the retrospective did not negatively influence students' accuracy in subsequent problem solving. Mainly, students' accuracy either improved or remained accurate. For comparison, when students did not receive an intervention, their accuracy showed little change—also when students were initially inaccurate (see Figure 12, no reflection).

At the same time, the results suggest that reflections of higher complexity—where students more deeply engaged with their visual behavior, conceptual understanding, or problem-solving strategies by evaluating alternatives or anticipating consequences—were associated with greater changes in their eye movements. In contrast, reflections of lower complexity, in which students engaged less deeply with these aspects, corresponded to smaller changes (see Figure 15, study I). Although intentions do not always translate directly into behavior (Sheeran & Webb, 2016), the findings indicate that students' intentions were mirrored in their eye-movement behavior.

Interestingly, when students did not engage in any reflection, they exhibited eye-movement changes similar in magnitude to those who reflected with higher complexity (see Figure 15, study II). Since these students did not engage with aspects of their problem solving in a reflection, such changes likely reflect a mixture of individual differences, stimulus-driven attention to salient features, or still-developing problem-solving expertise rather than purposeful adjustments (for further discussion, see Chapter Changes in Eye Movements).

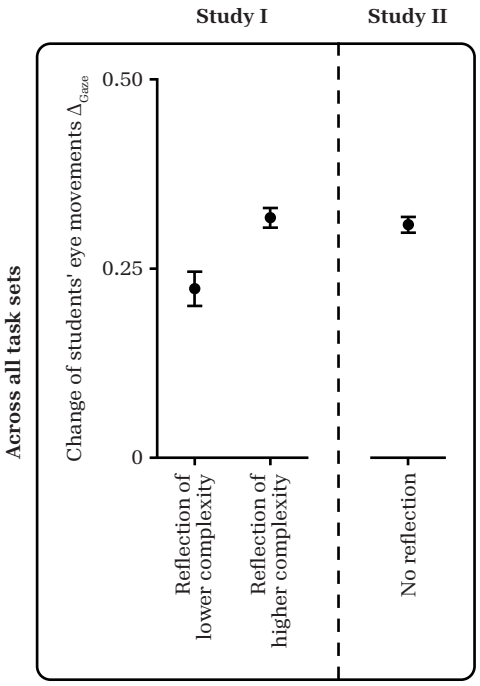


Figure 15 | Mean change of students' eye movements across all task sets in relation to students' reflection in study I (left). For comparison, the change of students' eye movements of study II are depicted (right).

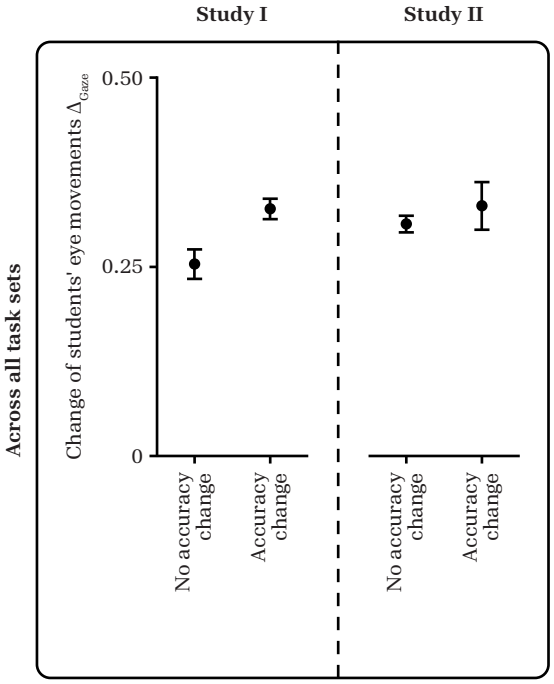


Figure 16 | Mean change of students' eye movements across all task set in relation to students' accuracy change in study I (left). For comparison, the change of students' eye movements of study II are depicted (right).

Given that visual attention filters information that cannot be processed simultaneously (Evans et al., 2011) and eye movements reflect underlying cognitive processes (Just & Carpenter, 1980; Lauwereyns & d'Ydewalle, 1996; Liu & Cui, 2025; Rayner, 1998), the retrospective appears to guide students toward a more intentional allocation of attention—promoting goal-driven processing and reducing stimulus- and history-driven processing (cf. Anderson, 2013; Awh et al., 2012; Theeuwes, 2019). Such a shift may contribute to strengthening students' representational competence by enabling them to more effectively interpret and use of representations (cf. Kozma & Russell, 2005; Rau, 2017).

Taken together, the findings indicate that the reflection processes stimulated during the eye-gaze-augmented retrospective influence students' performance—both in terms of accuracy and eye movements. Because accuracy and eye movements both reflect how students engage with and navigate a task, it is unsurprising that changes in accuracy are associated with changes in eye movements (see Figure 16, study I). This observation aligns with prior eye-tracking research demonstrating close links between eye movements and task accuracy (Baluyut & Holme, 2019; Braun et al., 2022; Karch et al., 2019; Pavlin et al., 2019; Pavlin & Slapničar, 2021; Slapničar et al., 2021; Slapničar et al., 2020; Tang et al., 2014; Weinrich & Britt, 2022). However, this trend is less pronounced in study II, as the variation in students' eye-movement changes among those who showed an accuracy change suggests that the groups are not as clearly distinguishable (see Figure 16, study II).

While accuracy and eye movements may appear to diverge in relation to reflection complexity, the findings suggest a

common underlying mechanism: whether students exhibit lower or higher reflection complexity, the retrospective is associated with more goal-driven allocation of attention, alongside either preserved or improved accuracy.

To examine whether the relationships between reflection complexity and changes in performance differ depending on the visual and conceptual affordances of organic chemistry tasks, students' changes in score and their eye movements were compared for each task context individually.

In the contexts R/S configuration of asymmetric carbon atoms and Brønsted acid-base concept, reflections of lower complexity were primarily shown by students who remained accurate across both tasks, whereas reflections of higher complexity were shown by students who either improved their score or remained accurate. In contrast, in the electrophilicity and nucleophilicity context, reflections of lower complexity were shown by students who either remained accurate or did not maintain accuracy, while reflections of higher complexity were mostly shown by students who improved their accuracy (see Figure 17, study I). This observation is grounded in the capability of students solving the initial problem rather than in differential influences of reflection complexity. Whereas most students were able to determine the structural site of tasks 1 in the R/S configuration of asymmetric carbon atoms and Brønsted acid-base concept contexts, almost no student could do this in the electrophilicity and nucleophilicity context. Identifying the β-carbon of the α,β-unsaturated carbonyl as an electrophilic center exceeded their conceptual knowledge. A similar observation emerged in study II (see Figure 17, study II), where most students were likewise unable to determine the structural site of tasks 1 in the electrophilicity and nucleophilicity context. Nevertheless, a trend appears across all contexts: students who engaged in higher reflection complexity during the retrospective either improved or remained accurate on the subsequent task, whereas students who did not receive a retrospective showed little change in accuracy.

Similarly, reflection complexity shows a similar influence on mean eye-movement changes in all three contexts in both studies: higher reflection complexity corresponded to larger eye-movement changes, lower reflection complexity to smaller ones, and no reflection were associated with changes comparable to those seen in higher reflections complexity (see Figure 18).

However, when examining the relationship between students' mean changes in eye movements and their changes in accuracy across contexts in both studies, the pattern of Figure 16 only emerges for the context R/S configuration of asymmetric carbon atoms (see Figure 19). In the Brønsted acid-base concept and electrophilicity and nucleophilicity contexts, eye-movement changes do not clearly differentiate between students whose accuracy changed and those whose accuracy did not change. These results suggest that changes in eye movements co-vary with changes in accuracy only in the R/S configuration of asymmetric carbon atoms context. A likely explanation is that the visual and conceptual affordances of the three contexts differ substantially.

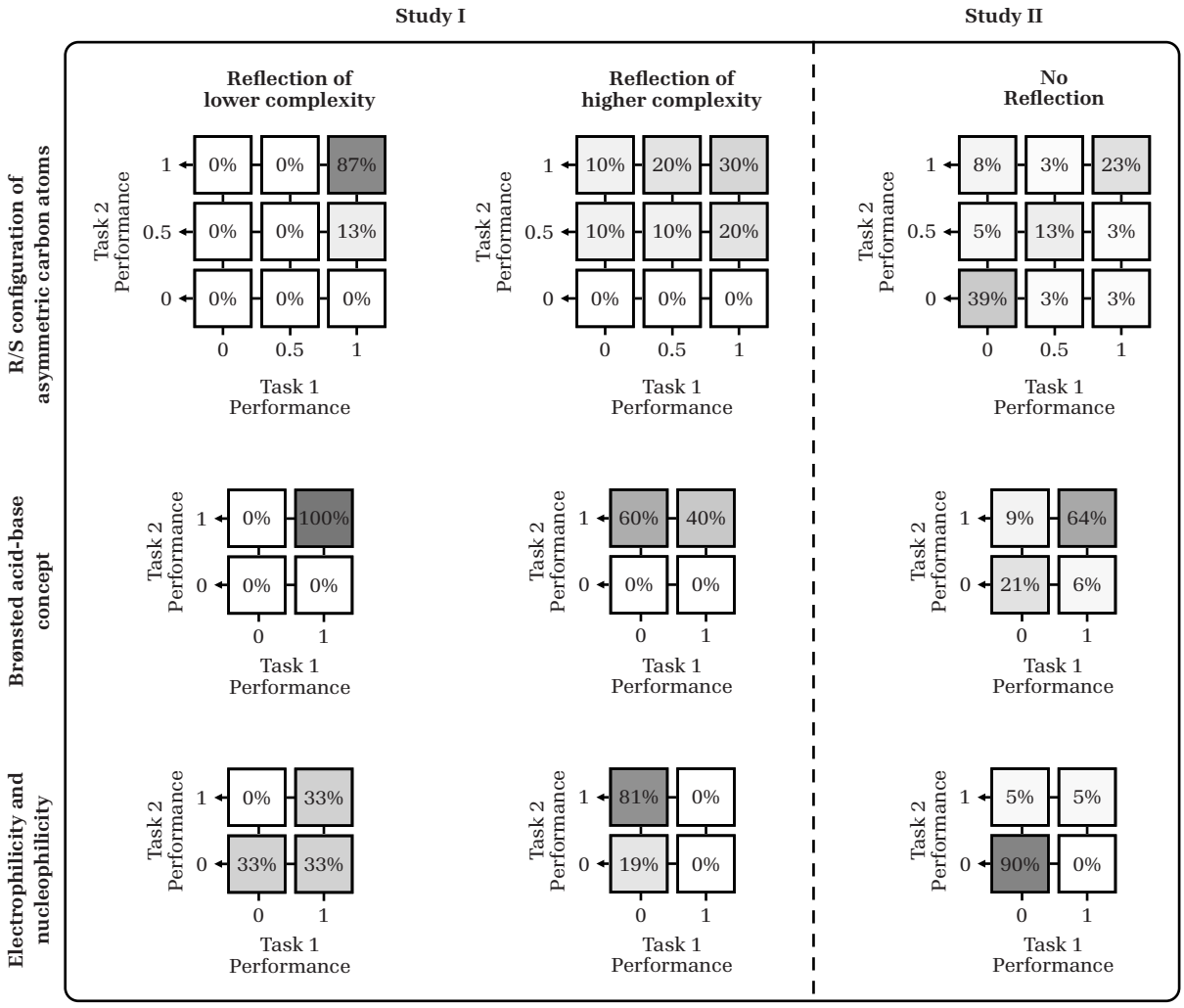


Figure 17 | Change of students' score in each task context in relation to reflection complexity in study I (left). Note: In study I, in reflection of lower complexity in the electrophilicity and nucleophilicity context the sample size was low (N = 3). For comparison, the change of students' score of study II are depicted (right).

The R/S configuration of asymmetric carbon atoms task requires students to visually inspect structural features in a systematic sequence—identifying the chiral carbon, assigning substituent priorities, and determining spatial orientation using established procedures (cf. Cahn et al., 1966; Dietzel, 1979; Idoux, 1982). This indicates higher visual affordances and comparatively lower conceptual affordances, suggesting that changes in students' problem-solving accuracy are more directly mirrored in their eye movements.

In contrast, in the Brønsted acid-base and electrophilicity and nucleophilicity tasks, once students perceive the relevant structural site, the task primarily requires reasoning with concepts such as electronegativity, resonance, or atomic size effects (cf. Fox & Whitesell, 2004; Klein, 2020). This involve higher conceptual affordances and comparatively lower visual affordances, shifting processing toward internal attention and resulting in less observable eye movement changes.

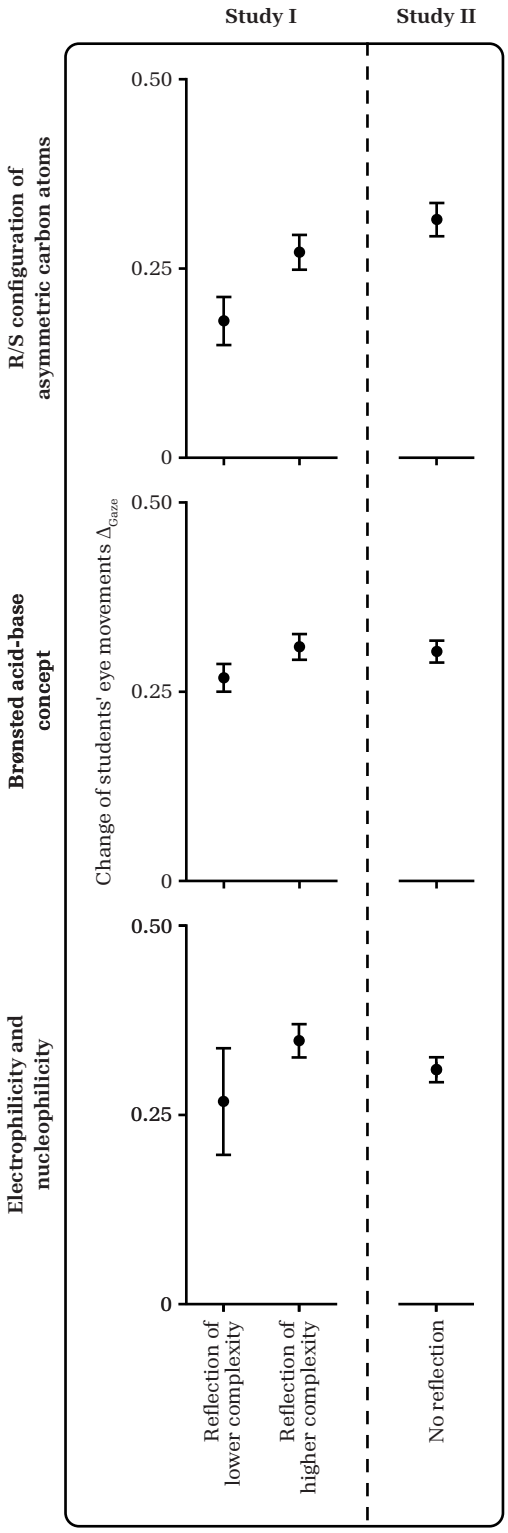


Figure 18 | Mean change of students' eye movements in each task pair in relation to students' reflection in study I (left). Note: In study I, in reflection of lower complexity in the electrophilicity and nucleophilicity task set the sample size was low ($N = 3$). For comparison, the change of students' eye movements of study II are depicted (right).

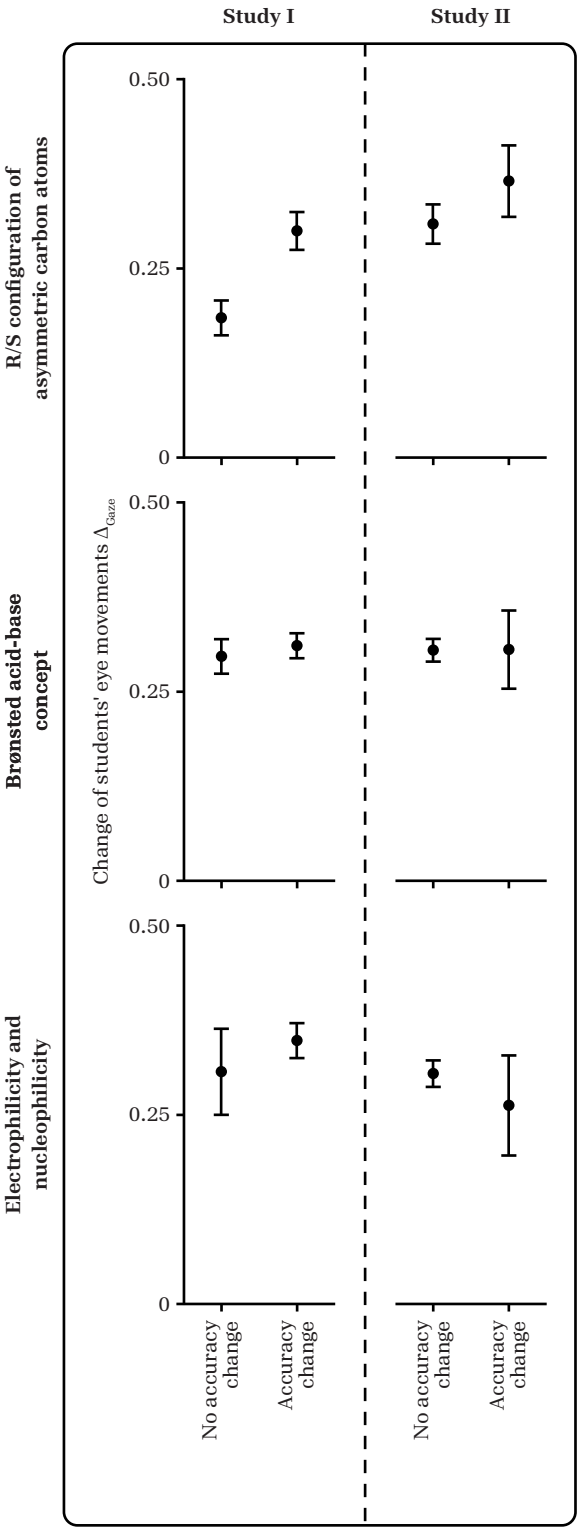


Figure 19 | Mean change of students' eye movements in each task pair in relation to students' performance change in study I (left). For comparison, the change of students' eye movements of study II are depicted (right). Note: In study II, in performance change in the electrophilicity and nucleophilicity task set the sample size was low ($N = 2$).

Taken together, these findings indicate that the influence of reflection complexity on students' subsequent problem-solving performance (i.e., accuracy and eye movements) does not differ across task contexts. However, the degree to which changes in accuracy co-vary with changes in eye movements appears greater in tasks with higher visual affordances and comparatively lower conceptual demands. In more conceptually demanding tasks, changes in accuracy are less reflected in shifts in eye movements.

BUT?

(i.e., Limitations and Research Implications)

Some limitations must be considered when interpreting the findings of this investigation (for further limitations, see Chapters Students' Reflection, Defining Areas of Interest, and Changes in Eye Movements)

First, this analysis examined reflection processes tied to specific key structural sites in isolation. However, due to the task design (see Figures 4) and retrospectives' structure (see Figures 5), students may have reflected on multiple sites concurrently. Such overlap may have interrupted the reflection before resuming in the next phase. Future studies could design tasks structured around one single key structural site, ensuring that participants reflect on one problem-solving process at a time.

Second, the operationalization of reflection complexity using four deductively derived codes provides a simplified characterization of students' reflection (see Tables 1 and 2). Given that existing reflection frameworks were developed in other domains, future inductive analyses may uncover additional facets of reflection specific to science problem solving not captured by the present analysis.

Third, studies I and II are not directly comparable. Although both samples consisted of Organic Chemistry 1 students, they differed in several aspects (e.g., native language, course instructor, textbook), and these differences must be considered when interpreting cross-study comparisons. While the patterns observed across the two studies may provide initial insights of the potential of gaze-augmented reflection, future research should verify these observations with directly comparable samples.

Fourth, although the sample sizes of both studies are typical for exploratory eye-tracking research in chemistry education (Havanki & VandenPlas, 2014), some subgroups were particularly small, limiting both the use of inferential statistics and the certainty of interpretations. Larger and more balanced samples in future work are needed to increase confidence in the effects of gaze-augmented reflection.

Finally, although the contexts examined here cover different visual and conceptual affordances, they do not represent the full range found in organic chemistry. Future studies should therefore explore additional contexts (e.g., hybridization, E/Z isomerism, aromaticity). Extending the intervention to other scientific domains may also be valuable, as students in many areas encounter difficulties similar to those observed in organic chemistry (cf. Ambrose et al., 1999; Bollen et al., 2016; Bollen et al., 2017; Bollen et al., 2015; Chi et al., 1981; Klein et al., 2018; Krajcik, 1991).

WHAT NOW?

(i.e., Conclusion)

This investigation examined the extent to which reflections stimulated by the eye-gaze-augmented retrospective shapes students' subsequent problem-solving performance. These findings allow us to draw a final conclusion about how gaze-augmented reflection shapes and supports students' problem-solving behavior.

The findings indicate that the complexity of students' reflection is related to, but not solely determined by, their initial problem-solving accuracy. Regardless of whether students exhibited lower or higher levels of reflection complexity, engaging in the eye-gaze-augmented retrospective was associated with more goal-driven allocation of visual attention during subsequent problem-solving, accompanied by either maintained or improved accuracy. Finally, the relationship between reflection complexity and changes in performance did not differ fundamentally across contexts with different visual and conceptual affordances—however, changes in accuracy were more clearly mirrored by changes in eye movements in tasks with higher visual and lower conceptual affordances, whereas in more conceptually demanding tasks changes in accuracy were less directly reflected in shifts of eye movements.

Beyond these findings observed across all students, the results further demonstrate that students show highly individual trajectories during the initial problem-solving, the eye-gaze-augmented retrospective, and the subsequent problem-solving. Depending on the task context, varying proportions of Organic Chemistry I students were able to solve the problems accurately. Partially based on their accuracy, students engaged in reflection in different ways—nevertheless, reflection complexity also varied as a function of individual student characteristics, such as metacognitive monitoring and regulation and achievement goals. Nevertheless, across these diverse trajectories, the outcome was largely similar: students' performance converged, with students either improving or maintaining their problem-solving performance.

Taken together, these findings suggest that gaze-augmented reflection holds considerable potential to support students' problem-solving with representations as a personalized learning experience.

WHAT NEXT?

(i.e., Outlook and Teaching Implications)

Based on the findings, the gaze-augmented reflection, by accommodating students' individual difficulties, differences, and learning consequences, demonstrates considerable potential for integration into an adaptive intelligent tutoring system for personalized learning based on artificial intelligence (AI). Such a system could support students individually along their own learning pathways (Lin et al., 2023). The application of AI in this context is versatile in two ways. First, students' task responses and reflections—provided in written or audio form—can be processed and categorized immediately (cf. Martin et al., 2024; Wulff et al., 2023). Second, by drawing on individual student characteristics (e.g., conceptual knowledge, cognitive reflection ability, achievement goals, and metacognitive monitoring and regulation) alongside data from their problem solving (e.g., accuracy, eye movements) and reflection (e.g., reflection complexity), the AI tutor could generate individualized prompts and provide targeted conceptual support when task demands exceed students' knowledge. In doing so, the system could guide each student through an eye-gaze-augmented retrospective, scaffold self-regulation in accordance with students' developmental stages (Panadero, 2017), and support students in a personalized manner as they identify and reflect on their experienced issues. A possible structure for such an adaptive learning system is illustrated in Figure 20.

As depicted in Figure 20, this adaptive learning environment integrates students' individual cognitive, motivational, and metacognitive characteristics with data from their problem-solving behavior to deliver personalized guidance during an eye-gaze-augmented retrospective. At the outset, students complete a set of assessments capturing their conceptual knowledge (e.g., Farhat et al., 2019; Nedungadi et al., 2021; Steinbach et al., 2025), cognitive reflection abilities (e.g., Frederick, 2005; Toplak et al., 2014), achievement goals (e.g., Elliot & McGregor, 2001), and metacognitive monitoring and regulation (e.g., Cooper et al., 2008; Schraw & Dennison, 1994) (see Figure 20, tests). These automatically processed measures form the foundation for tailoring subsequent support to each student's individual needs (see Figure 20, adaption).

Following this diagnostic phase, students engage in an initial problem-solving task (see Figure 20, initial task). During this task, their eye movements are recorded and their problem-solving accuracy is logged as indicators of task performance (cf. Dinc et al., 2025) (see Figure 20, adaption).

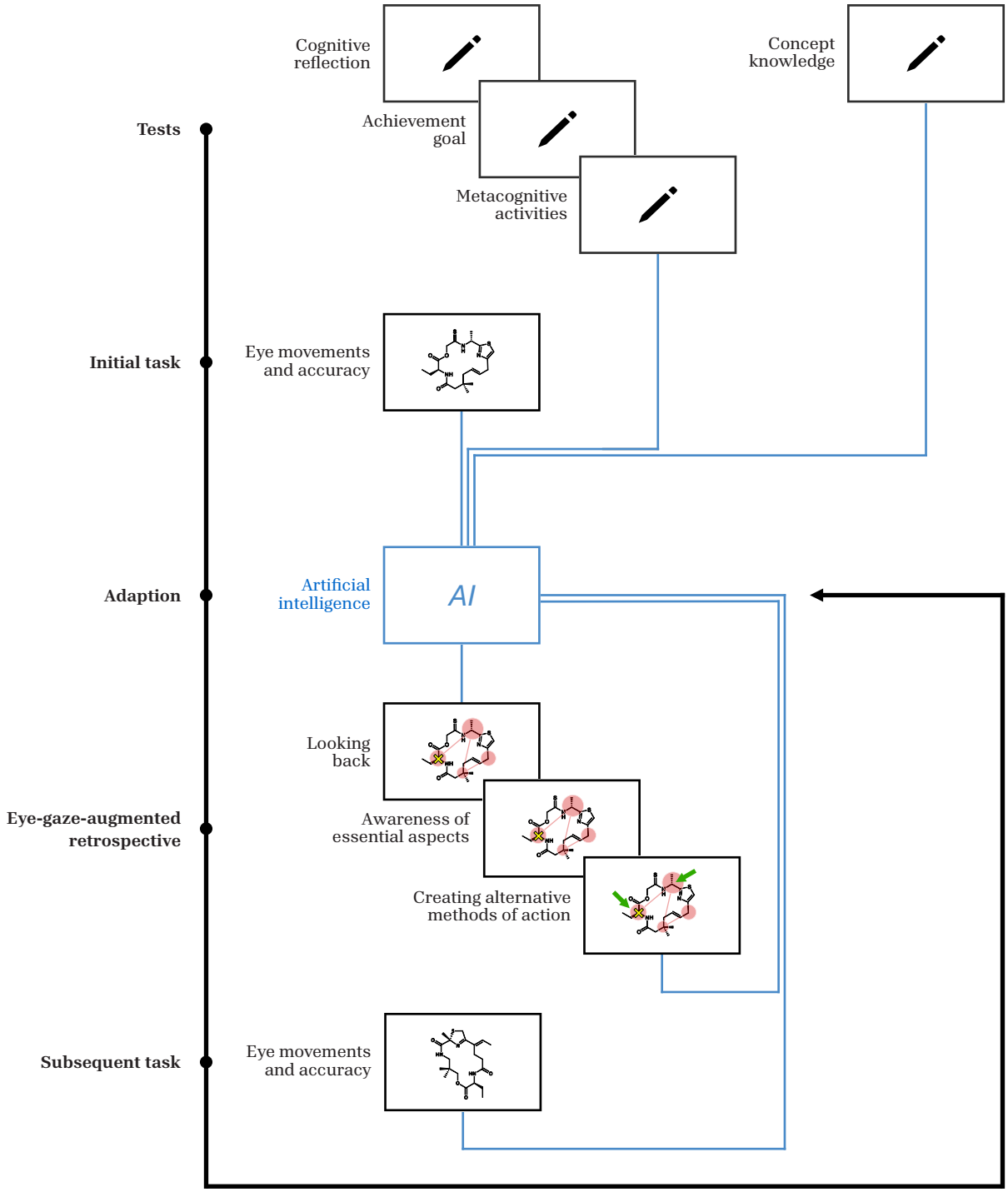


Figure 20 | Structure of the intelligent tutoring system (black) based on artificial intelligence (blue).

Students then enter an eye-gaze-augmented retrospective (see Figure 20, eye-gaze-augmented retrospective). Drawing on information about students' cognitive reflection ability, achievement goal orientation, metacognitive tendencies, eye movements, and problem-solving accuracy, the tutor provides individualized reflection prompts designed to scaffold a structured examination of their problem-solving processes. In addition, students receive conceptual support aligned with their performance on the conceptual knowledge assessment, ensuring that reflection is not constrained by gaps in conceptual understanding. At the procedural level, the AI tutor may also provide adaptive prompts—for example, following the Goldilocks Help (Yuriev et al., 2017)—to support students when they encounter dead ends or false starts during reflection. Within this guided retrospective, the AI functions as an adaptive conversational agent that dynamically responds to students' reflections and prompts them to describe their experiences, identify issues, explore alternative approaches, and draw consequences for future problem solving. During this process, the system evaluates the complexity of students' reflections (see Chapter Students' Reflection).

Students then engage in a subsequent, structurally similar problem-solving task (see Figure 20, subsequent task), during which eye movements and task accuracy are again recorded. The system can then examine changes in eye movements (see Chapters Defining Areas of Interest and Changes in Eye Movements) in relation to the complexity of students' reflections, enabling the detection of shifts toward more goal-driven visual processing.

From this point onward, the learning setup can continue in iterative cycles of problem-solving, AI processing, and eye-gaze-augmented retrospective, allowing students to receive support for as long as they choose. Across successive cycles, the AI tutor can provide additional prompts or visual scaffolds when students exhibit tendencies toward stimulus-driven visual behavior.

Such an adaptive learning system could accompany an organic chemistry course throughout a semester, allowing students to voluntarily and dynamically adjust their training according to their individual learning progress and needs.

Overall, such an intelligent tutoring system illustrates how the integration of multimodal data with adaptive scaffolding could create a personalized pathway that supports students' conceptual, procedural, and visual processing of organic chemistry problems with representations, thereby addressing students' challenges that have been documented in chemistry education research (as reviewed by Graulich, 2015b, 2025).

.

REFERENCES

(This time, the title is academically appropriate.)

Ambrose, B. S., Heron, P. R. L., Vokos, S., & McDermott, L. C. (1999). Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena. *American Journal of Physics*, *67*(10), 891–898. https://doi.org/10.1119/1.19144

Anderson, B. A. (2013). A value-driven mechanism of attentional selection. *Journal of Vision*, 13(3), 1–16. https://doi.org/10.1167/13.3.7

Anzovino, M. E., & Bretz, S. L. (2015). Organic chemistry students' ideas about nucleophiles and electrophiles: the role of charges and mechanisms. *Chemistry Education Research and Practice*, *16*(4), 797–810. https://doi.org/10.1039/C5RP00113G

Anzovino, M. E., & Bretz, S. L. (2016). Organic chemistry students' fragmented ideas about the structure and function of nucleophiles and electrophiles: a concept map analysis. *Chemistry Education Research and Practice*, *17*(4), 1019–1029. https://doi.org/10.1039/C6RP00111D

Arsic, B., Novak, P., Kragol, G., Barber, J., Rimoli, M. G., & Sodano, F. (2018). *Macrolides*. De Gruyter. https://doi.org/10.1515/9783110515756

Atkins, S., & Murphy, K. (1993). Reflection: a review of the literature. *Journal of Advanced Nursing*, *18*(8), 1188–1192. https://doi.org/10.1046/j.1365-2648.1993.18081188.x

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*(8), 437–443. https://doi.org/10.1016/j.tics.2012.06.010

Baluyut, J. Y., & Holme, T. A. (2019). Eye tracking student strategies for solving stoichiometry problems involving particulate nature of matter diagrams. *Chemistry Teacher International*, *1*(1), 1–13. https://doi.org/10.1515/cti-2018-0003

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The Effects of School-Based Writing-to-Learn Interventions on Academic Achievement: A Meta-Analysis. *Review of Educational Research*, *74*(1), 29–58. https://doi.org/10.3102/00346543074001029

Bhattacharyya, G., & Bodner, G. M. (2005). "It Gets Me to the Product": How Students Propose Organic Mechanisms. *Journal of Chemical Education*, *82*(9), 1402. https://doi.org/10.1021/ed082p1402

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. Annual Review of Psychology, 64(Volume 64, 2013), 417–444. https://doi.org/10.1146/annurev-psych-113011-143823

Bollen, L., van Kampen, P., Baily, C., & De Cock, M. (2016). Qualitative investigation into students' use of divergence and curl in electromagnetism. *Physical Review Physics Education Research*, *12*(2), 020134. https://doi.org/10.1103/PhysRevPhysEducRes.12.020134

Bollen, L., van Kampen, P., Baily, C., Kelly, M., & De Cock, M. (2017). Student difficulties regarding symbolic and graphical representations of vector fields. *Physical Review Physics Education Research*, *13*(2), 020109. https://doi.org/10.1103/PhysRevPhysEducRes.13.020109

Bollen, L., van Kampen, P., & De Cock, M. (2015). Students' difficulties with vector calculus in electrodynamics. Physical Review Special Topics - *Physics Education Research*, *11*(2), 020129. https://doi.org/10.1103/PhysRevSTPER.11.020129

Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. Routledge.

Boyd, E. M., & Fales, A. W. (1983). Reflective Learning: Key to Learning from Experience. *Journal of Humanistic Psychology*, *23*(2), 99–117. https://doi.org/10.1177/0022167883232011

Brandriet, A. R., & Bretz, S. L. (2014). Measuring meta-ignorance through the lens of confidence: examining students' redox misconceptions about oxidation numbers, charge, and electron transfer. *Chemistry Education Research and Practice*, *15*(4), 729–746. https://doi.org/10.1039/C4RP00129J

Bratfisch, O., Borg, G., & Dornič, S. (1972). *Perceived Item-difficulty in Three Tests of Intellectual Performance Capacity*.

Braun, I., & Graulich, N. (2024a). Exploring diversity: student's (un-)productive use of resonance in organic chemistry tasks through the lens of the coordination class theory. *Chemistry Education Research and Practice*, *25*(3), 643–671. https://doi.org/10.1039/D3RP00298E

Braun, I., & Graulich, N. (2024b). The More Explicit, the Better? How the Indication of Electron Lone Pairs Affects Students' Consideration of Resonance in Organic Chemistry. *Journal of Chemical Education*, *101*(11), 4830–4836. https://doi.org/10.1021/acs.jchemed.4c00762

Braun, I., Langner, A., & Graulich, N. (2022). Let's draw molecules: Students' sequential drawing processes of resonance structures in organic chemistry. *Frontiers in Education*, *7*, 1–24. https://doi.org/10.3389/feduc.2022.1055280

Braun, I., Lewis, S. E., & Graulich, N. (2025). A question of pattern recognition: investigating the impact of structure variation on students' proficiency in deciding about resonance stabilization. *Chemistry Education Research and Practice*, *26*(1), 158–182. https://doi.org/10.1039/D4RP00155A

Buddrus, J., & Schmidt, B. (2015). *Grundlagen der Organischen Chemie*. De Gruyter. https://doi.org/10.1515/9783110331059

Cahn, R. S., Ingold, C., & Prelog, V. (1966). Specification of Molecular Chirality. *Angewandte Chemie International Edition in English*, *5*(4), 385–415. https://doi.org/10.1002/anie.196603851

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*(2), 121–152.

Christian, K., & Talanquer, V. (2012). Modes of reasoning in self-initiated study groups in chemistry. *Chemistry Education Research and Practice*, *13*(3), 286–295. https://doi.org/10.1039/C2RP20010D

Conroy, D. E., Elliot, A. J., & Thrash, T. M. (2009). Achievement motivation. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of Individual Differences in Social Behavior* (pp. 382–399). The Guilford Press.

Cooper, M. M., Corley, L. M., & Underwood, S. M. (2013). An investigation of college chemistry students' understanding of structure–property relationships. *Journal of Research in Science Teaching*, *50*(6), 699–721. https://doi.org/10.1002/tea.21093

Cooper, M. M., Sandi–Urena, S., & Stevens, R. (2008). Reliable multi method assessment of metacognition use in chemistry problem solving. *Chemistry Education Research and Practice*, *9*(1), 18–24. https://doi.org/10.1039/B801287N

Crowder, C. J., & Raker, J. R. (2024). Patterns in Explanations of Organic Chemistry Reaction Mechanisms: A Text Analysis by Level of Explanation Sophistication. *Journal of Chemical Education*, *101*(12), 5203–5220. https://doi.org/10.1021/acs.jchemed.4c01042

Cruz-Ramírez de Arellano, D., & Towns, M. H. (2014). Students' understanding of alkyl halide reactions in undergraduate organic chemistry. *Chemistry Education Research and Practice*, *15*(4), 501–515. https://doi.org/10.1039/C3RP00089C

Cullipher, S., & Sevian, H. (2015). Atoms versus Bonds: How Students Look at Spectra. *Journal of Chemical Education*, *92*(12), 1996–2005. https://doi.org/10.1021/acs.jchemed.5b00529

Davis, E. A. (2003). Prompting Middle School Science Students for Productive Reflection: Generic and Directed Prompts. *The Journal of the Learning Sciences*, *12*(1), 91–142. https://doi.org/10.1207/S15327809JLS1201_4

DeFever, R. S., Bruce, H., & Bhattacharyya, G. (2015). Mental rolodexing: Senior chemistry majors' understanding of chemical and physical properties. *Journal of Chemical Education*, *92*(3), 415–426. https://doi.org/10.1021/ed500360g

Dewey, J. (1933). *How we think*. Heath & Co Publishers.

Dietzel, R. A. (1979). Determination of chiral molecule configuration using the ±1,2,5 Rule. *Journal of Chemical Education*, *56*(7), 451. https://doi.org/10.1021/ed056p451

Dinc, Y., Malone, S., Ruf, V., Steinert, S., Küchemann, S., & Kuhn, J. (2025). Gaze-based prediction of test performance using machine learning for adaptive learning systems. *Physical Review Physics Education Research*, *21*(2), 020148. https://doi.org/10.1103/4yys-22tc

Domin, D. S., Al-Masum, M., & Mensah, J. (2008). Students' categorizations of organic compounds. *Chemistry Education Research and Practice*, *9*(2), 114–121. https://doi.org/10.1039/B806226A

Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice* (3. ed.). Springer. https://doi.org/10.1007/978-3-319-57883-5

Dumont, H., & Ready, D. D. (2023). On the promise of personalized learning for educational equity. *npj Science of Learning*, *8*(1), 26. https://doi.org/10.1038/s41539-023-00174-x

Elby, A. (2000). What students' learning of representations tells us about constructivism. *The Journal of Mathematical Behavior*, *19*(4), 481–502. https://doi.org/10.1016/S0732-3123(01)00054-2

Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. The MIT Press.

Evans, K. K., Horowitz, T. S., Howe, P., Pedersini, R., Reijnen, E., Pinto, Y., Kuzmova, Y., & Wolfe, J. M. (2011). Visual attention. *WIREs Cognitive Science*, *2*(5), 503–514. https://doi.org/10.1002/wcs.127

Farhat, N. J., Stanford, C., & Ruder, S. M. (2019). Assessment of Student Performance on Core Concepts in Organic Chemistry. *Journal of Chemical Education*, *96*(5), 865–872. https://doi.org/10.1021/acs.jchemed.8b00913

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human perception and performance*, *1*(3), 288–299.

Fleck, R., & Fitzpatrick, G. (2010). *Reflecting on reflection: framing a design landscape*. OZCHI '10, Brisbane, Australia.

Fox, M. A., & Whitesell, J. K. (2004). *Organic Chemistry* (3. ed.). Jones and Bartlett Publishers.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, *19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Frey, R. F., Brame, C. J., Fink, A., & Lemons, P. P. (2022). Teaching Discipline-Based Problem Solving. *CBE—Life Sciences Education*, *21*(2), 21:fe21. https://doi.org/10.1187/cbe.22-02-0030

Galloway, K. R., Leung, M. W., & Flynn, A. B. (2018). A Comparison of How Undergraduates, Graduate Students, and Professors Organize Organic Chemistry Reactions. *Journal of Chemical Education*, *95*(3), 355–365. https://doi.org/10.1021/acs.jchemed.7b00743

Garrett, R. M. (1986). Problem-solving in Science Education. *Studies in Science Education*, *13*(1), 70–95. https://doi.org/10.1080/03057268608559931

Garrett, R. M. (1987). Issues in science education: problem-solving, creativity and originality. *International Journal of Science Education*, *9*(2), 125–137. https://doi.org/10.1080/0950069870090201

Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Oxford Brookes University.

Glenberg, A. M., Wilkinson, A. C., & Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory & Cognition*, *10*(6), 597–602. https://doi.org/10.3758/BF03202442

Goodwin, W. (2010). How do Structural Formulas Embody the Theory of Organic Chemistry? *The British Journal for the Philosophy of Science*, *61*(3), 621–633. https://doi.org/10.1093/bjps/axp052

Graulich, N. (2015a). Intuitive Judgments Govern Students' Answering Patterns in Multiple-Choice Exercises in Organic Chemistry. *Journal of Chemical Education*, *92*(2), 205–211. https://doi.org/10.1021/ed500641n

Graulich, N. (2015b). The tip of the iceberg in organic chemistry classes: how do students deal with the invisible?. *Chemistry Education Research and Practice*, *16*(1), 9–21. https://doi.org/10.1039/C4RP00165F

Graulich, N. (2025). The tip of the iceberg in organic chemistry – revisited. *Chemistry Education Research and Practice*, *26*(2), 359–376. https://doi.org/10.1039/D4RP00345D

Graulich, N., & Bhattacharyya, G. (2017). Investigating students' similarity judgments in organic chemistry. *Chemistry Education Research and Practice*, *18*(4), 774–784. https://doi.org/10.1039/C7RP00055C

Graulich, N., Hedtrich, S., & Harzenetter, R. (2019). Explicit versus implicit similarity – exploring relational conceptual understanding in organic chemistry. *Chemistry Education Research and Practice*, *20*(4), 924–936. https://doi.org/10.1039/C9RP00054B

Guo, L. (2022). How should reflection be supported in higher education? — A meta-analysis of reflection interventions. *Reflective Practice*, *23*(1), 118–146. https://doi.org/10.1080/14623943.2021.1995856

Habermas, J. (1974). *Theory and Practice* (J. Viertel, Trans.; 4. ed.). Beacon Press.

Hamel, C., & Viau-Guay, A. (2019). Using video to support teachers' reflective practice : A literature review. *Cogent Education*, *6*(1), 1673689. https://doi.org/10.1080/2331186X.2019.1673689

Hansen, S. J. R., Hu, B., Riedlova, D., Kelly, R. M., Akaygun, S., & Villalta-Cerdas, A. (2019). Critical consumption of chemistry visuals: eye tracking structured variation and visual feedback of redox and precipitation reactions. *Chemistry Education Research and Practice*, *20*(4), 837–850. https://doi.org/10.1039/c9rp00015a

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hartmann, C., van Gog, T., & Rummel, N. (2021). Preparatory effects of problem solving versus studying examples prior to instruction. *Instructional Science*, *49*(1), 1–21. https://doi.org/10.1007/s11251-020-09528-z

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. https://doi.org/10.1016/0742-051X(94)00012-U

Havanki, K. L., & VandenPlas, J. R. (2014). Eye Tracking Methodology for Chemistry Education Research. In D. M. Bunce & R. S. Cole (Eds.), *Tools of Chemistry Education Research* (Vol. 1166, pp. 191–218). American Chemical Society. https://doi.org/10.1021/bk-2014-1166.ch011

Hinze, S. R., Rapp, D. N., Williamson, V. M., Shultz, M. J., Deslongchamps, G., & Williamson, K. C. (2013). Beyond ball-and-stick: Students' processing of novel STEM visualizations. *Learning and Instruction*, *26*, 12–21. https://doi.org/10.1016/j.learninstruc.2012.12.002

Hoffmann, R., & Laszlo, P. (1991). Representation in Chemistry. *Angewandte Chemie International Edition in English*, *30*(1), 1–16. https://doi.org/10.1002/anie.199100013

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J., & Lehtinen, M. (2008). Gaze path stimulation in retrospective think-aloud. *Journal of Eye Movement Research*, *2*(4), 1–18. https://doi.org/10.16910/jemr.2.4.5

Idoux, J. P. (1982). A simple method for specifying the R/S configuration about a chiral center. *Journal of Chemical Education*, *59*(7), 553. https://doi.org/10.1021/ed059p553

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review, 87*(4), 329–354.

Karch, J. M., García Valles, J. C., & Sevian, H. (2019). Looking into the black box: Using gaze and pupillometric data to probe how cognitive load changes with mental tasks. *Journal of Chemical Education*, *96*(5), 830–840. https://doi.org/10.1021/acs.jchemed.9b00014

Kember, D. (1999). Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *International Journal of Lifelong Education*, *18*(1), 18–30. https://doi.org/10.1080/026013799293928

Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K., Tse, H., Webb, C., Yuet Wong, F. K., Wong, M., & Yeung, E. (2000). Development of a Questionnaire to Measure the Level of Reflective Thinking. *Assessment & Evaluation in Higher Education*, *25*(4), 381–395. https://doi.org/10.1080/713611442

Killion, J. P., & Todnem, G. R. (1991). A Process for Personal Theory Building. *Educational Leadership*, *48*(6), 14–16.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, *31*(2), 338–368.

Klein, D. R. (2020). *Organic Chemistry* (4., Ed.). Wiley.

Klein, P., Viiri, J., Mozaffari, S., Dengel, A., & Kuhn, J. (2018). Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots? *Physical Review Physics Education Research*, *14*(1), 010116. https://doi.org/10.1103/PhysRevPhysEducRes.14.010116

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and Validation of Two Instruments Measuring Intrinsic, Extraneous, and Germane Cognitive Load. *Frontiers in Psychology*, *8*(1997), 1–18. https://doi.org/10.3389/fpsyg.2017.01997

Koh, D. H., Xu, Z., Wang, J., Burgess, A. N., Seccia, A., Schneps, M., Pomplun, M., Lamb, R., Keil, A., Dawson, K., & Antonenko, P. (2024). Fixation Disparity: A Possible Index of Visuospatial Cognition during Authentic Learning Tasks. *Mind, Brain, and Education*, *18*(3), 346–359. https://doi.org/10.1111/mbe.12424

Koole, S., Dornan, T., Aper, L., Scherpbier, A., Valcke, M., Cohen-Schotanus, J., & Derese, A. (2011). Factors confounding the assessment of reflection: a critical review. *BMC Medical Education*, *11*, 104. https://doi.org/10.1186/1472-6920-11-104

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review*, *11*, 45–55. https://doi.org/10.1016/j.edurev.2013.11.003

Korthagen, F. A. (1985). Reflective Teaching and Preservice Teacher Education in the Netherlands. *Journal of Teacher Education*, *36*(5), 11–15. https://doi.org/10.1177/002248718503600502

Korthagen, F. A., & Kessels, J. P. (1999). Linking Theory and Practice: Changing the Pedagogy of Teacher Education. *Educational Researcher*, *28*(4), 4–17. https://doi.org/10.3102/0013189X028004004

Kozma, R. B., & Russell, J. (2005). Students Becoming Chemists: Developing Representationl Competence. In J. K. Gilbert (Ed.), *Visualization in Science Education* (pp. 121–145). Springer Netherlands.

Krajcik, J. S. (1991). Developing students' understanding of chemical concepts. In S. M. Glynn, R. H. Yeany, & B. K. Britton (Eds.), *The Psychology of Learning Science* (pp. 117–148). Routledge.

Kranz, D., Martin, P. P., Schween, M., & Graulich, N. (2025). Should we scaffold it? Analysing the effect of task format and scaffolding on students' learning gain. Chemistry Education Research and Practice. https://doi.org/10.1039/D4RP00241E

Kranz, D., Schween, M., & Graulich, N. (2023). Patterns of reasoning – exploring the interplay of students' work with a scaffold and their conceptual knowledge in organic chemistry. *Chemistry Education Research and Practice*, *24*(2), 453–477. https://doi.org/10.1039/D2RP00132B

Kruglanski, A. W., & Webster, D. M. (1996). Motivated Closing of the Mind: "Seizing" and "Freezing.". *Psychological review*, *103*(2), 263–283. https://doi.org/10.1037/0033-295X.103.2.263

Langner, A., & Graulich, N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804

Langner, A., Hain, L., & Graulich, N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, *102*(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830

Langner, A., Sahba, M., Popova, M., & Graulich, N. (2025). An Integrated Approach to Characterizing Changes in Organic Chemistry Students' Eye Movements. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

Lauwereyns, J., & d'Ydewalle, G. (1996). Knowledge acquisition in poetry criticism: the expert's eye movements as an information tool. *International Journal of Human-Computer Studies*, *45*(1), 1–18. https://doi.org/10.1006/ijhc.1996.0039

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Lin, C.-C., Huang, A. Y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learning Environments*, *10*(1), 41. https://doi.org/10.1186/s40561-023-00260-y

Lin, H.-S., Hung, J.-Y., & Hung, S.-C. (2002). Using the history of science to promote students' problem-solving ability. *International Journal of Science Education*, *24*(5), 453–464. https://doi.org/10.1080/09500690110073991

Liu, X., & Cui, Y. (2025). Eye tracking technology for examining cognitive processes in education: a systematic review. *Computers & Education*, *229*, 105263. https://doi.org/10.1016/j.compedu.2025.105263

Locatelli, S. W., de Oliveira, L. S., da Silva S. Junior, R., & Sato, J. R. (2024). Perspectives of the Eye-Tracking Application to Understand Access to Representational Levels in Chemistry. *Journal of Science Education and Technology*, *33*, 823–834. https://doi.org/10.1007/s10956-024-10123-x

Marshall, T., Keville, S., Cain, A., & Adler, J. R. (2021). On being open-minded, wholehearted, and responsible: a review and synthesis exploring factors enabling practitioner development in reflective practice. *Reflective Practice*, *22*(6), 860–876. https://doi.org/10.1080/14623943.2021.1976131

Martin, P. P., Kranz, D., Wulff, P., & Graulich, N. (2024). Exploring new depths: Applying machine learning for the analysis of student argumentation in chemistry. *Journal of Research in Science Teaching*, *61*(8), 1757–1792. https://doi.org/doi.org/10.1002/tea.21903

Martinez, M. E. (1998). What Is Problem Solving? *The Phi Delta Kappan*, *79*(8), 605–609.

McClary, L., & Talanquer, V. (2011). Heuristic Reasoning in Chemistry: Making decisions about acid strength. *International Journal of Science Education*, *33*(10), 1433–1454. https://doi.org/10.1080/09500693.2010.528463

Meece, J. L. (1994). The role of motivation in self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 25–44). Routledge.

Mezirow, J. (1981). A Critical Theory of Adult Learning and Education. *Adult Education*, *32*(1), 3–24. https://doi.org/10.1177/074171368103200101

Mezirow, J. (1990). How critical reflection triggers transformative learning. In J. Mezirow (Ed.), *Fostering critical reflection in adulthood: A guide to transformative and emancipatory learning* (pp. 1–20). Jossey-Bass.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey-Bass.

Moon, J. A. (2004a). *A Handbook of Reflective and Experiential Learning: Theory and Practice*. Psychology Press.

Moon, J. A. (2004b). *Reflection in learning and professional development: Theory and practice*. Routledge.

Nedungadi, S., Mosher, M. D., Paek, S. H., Hyslop, R. M., & Brown, C. E. (2021). Development and psychometric analysis of an inventory of fundamental concepts for understanding organic reaction mechanisms. *Chemistry Teacher International*, *3*(4), 377–390. https://doi.org/10.1515/cti-2021-0009

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176–1189. https://doi.org/10.1111/medu.12583

Nicholls, J. G., & Miller, A. T. (1984). The Differentiation of the Concepts of Difficulty and Ability. In J. G. Nicholls (Ed.), *Advances in motivation and achievement* (3. ed., pp. 185–218). JAI Press.

Nückles, M., Roelle, J., Glogger-Frey, I., Waldeyer, J., & Renkl, A. (2020). The Self-Regulation-View in Writing-to-Learn: Using Journal Writing to Optimize Cognitive Load in Self-Regulated Learning. *Educational Psychology Review*, *32*(4), 1089–1126. https://doi.org/10.1007/s10648-020-09541-1

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00422

Pavlin, J., Glažar, S. A., Slapničar, M., & Devetak, I. (2019). The impact of students' educational background, interest in learning, formal reasoning and visualisation abilities on gas context-based exercises achievements with submicro-animations. *Chemistry Education Research and Practice*, *20*(3), 633–649. https://doi.org/10.1039/C8RP00189H

Pavlin, J., & Slapničar, M. (2021). The Impact of Students' Educational Background, Formal Reasoning, Visualisation Abilities, and Perception of Difficulty on Eye-Tracking Measures When Solving a Context-Based Problem with Submicroscopic Representation. In I. Devetak & S. A. Glažar (Eds.), *Applying Bio-Measurements Methodologies in Science Education Research* (pp. 217–241). Springer International Publishing. https://doi.org/10.1007/978-3-030-71535-9_11

Peltier, J. W., Hay, A., & Drago, W. (2005). The Reflective Learning Continuum: Reflecting on Reflection. *Journal of marketing education*, *27*(3), 250–263. https://doi.org/10.1177/027347530527965

Popova, M., & Bretz, S. L. (2018a). Organic chemistry students' challenges with coherence formation between reactions and reaction coordinate diagrams. *Chemistry Education Research and Practice*, *19*(3), 732–745. https://doi.org/10.1039/C8RP00064F

Popova, M., & Bretz, S. L. (2018b). Organic Chemistry Students' Understandings of What Makes a Good Leaving Group. *Journal of Chemical Education*, *95*(7), 1094–1101. https://doi.org/10.1021/acs.jchemed.8b00198

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211–227. https://doi.org/10.1002/sce.3730660207

Rau, M. A. (2017). Conditions for the Effectiveness of Multiple Visual Representations in Enhancing STEM Learning. *Educational Psychology Review, 29*(4), 717–761. https://doi.org/10.1007/s10648-016-9365-3

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction*, *16*(2), 92–103. https://doi.org/10.1016/j.learninstruc.2006.02.008

Rodemer, M., Karch, J., & Bernholt, S. (2023). Pupil dilation as cognitive load measure in instructional videos on complex chemical representations. *Frontiers in Education*, *8*, 1–11. https://doi.org/10.3389/feduc.2023.1062053

Rotich, F., Ward, L., Beck, C., & Popova, M. (2024). Attention is currency: how surface features of Lewis structures influence organic chemistry student reasoning about stability. *Chemistry Education Research and Practice*, *25*(4), 1071–1089. https://doi.org/10.1039/D4RP00030G

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: an illusion of explanatory depth. *Cognitive Science*, *26*(5), 521–562. https://doi.org/10.1207/s15516709cog2605_1

Rushton, G. T., Hardy, R. C., Gwaltney, K. P., & Lewis, S. E. (2008). Alternative conceptions of organic chemistry topics among fourth year chemistry students. *Chemistry Education Research and Practice*, *9*(2), 122–130. https://doi.org/10.1039/B806228P

Sandars, J. (2009). The use of reflection in medical education: AMEE Guide No. 44. *Medical Teacher*, *31*(8), 685–695. https://doi.org/10.1080/01421590903050374

Sandi-Urena, S., Loría Cambronero, G., & Jinesta Chaves, D. (2019). Conceptualisation of Lewis structures by chemistry majors. *Chemistry Teacher International*, *2*(1), 20180019. https://doi.org/10.1515/cti-2018-0019

Schön, D. A. (1983). The Reflective Practitioner. Temple Smith.

Schraw, G., & Dennison, R. S. (1994). Assessing Metacognitive Awareness. *Contemporary Educational Psychology*, *19*(4), 460–475. https://doi.org/10.1006/ceps.1994.1033

Sheeran, P., & Webb, T. L. (2016). The Intention–Behavior Gap. *Social and Personality Psychology Compass*, *10*(9), 503–518. https://doi.org/10.1111/spc3.12265

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education*, *8*(4), 485–503. https://doi.org/10.1007/s40692-021-00188-7

Sinha, T., & Kapur, M. (2021). When Problem Solving Followed by Instruction Works: Evidence for Productive Failure. *Review of Educational Research*, *91*(5), 761–798. https://doi.org/10.3102/00346543211019105

Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N. S., Tripoliti, E., Marias, K., Fotiadis, D. I., & Tsiknakis, M. (2021). Review of Eye Tracking Metrics Involved in Emotional and Cognitive Processes. *IEEE Reviews in Biomedical Engineering*, *16*, 260–277. https://doi.org/10.1109/RBME.2021.3066072

Slapničar, M., Tompa, V., Devetak, I., Glažar, S. A., & Pavlin, J. (2021). Using an Eye-Tracker to Study Students' Attention Allocation When Solving a Context-Based Problem on the Sublimation of Water. In I. Devetak & S. A. Glažar (Eds.), *Applying Bio-Measurements Methodologies in Science Education Research* (pp. 107–127). Springer International Publishing. https://doi.org/10.1007/978-3-030-71535-9_6

Slapničar, M., Tompa, V., Glažar, S. A., Devetak, I., & Pavlin, J. (2020). Students' Achievements in Solving Authentic Tasks with 3D Dynamic Sub-Microscopic Animations About Specific States of Water and their Transition. *Acta Chimica Slovenica*, *67*(3), 904–915. https://doi.org/10.17344/acsi.2020.5908

Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 15–38). Cambridge University Press.

Stanford, C., Moon, A., Towns, M., & Cole, R. (2018). The Impact of Guided Inquiry Materials on Student Representational Level Understanding of Thermodynamics. In C. M. Teague & D. E. Gardner (Eds.), Engaging Students in Physical Chemistry (Vol. 1279, pp. 141–168). *American Chemical Society*. https://doi.org/10.1021/bk-2018-1279.ch010

Steinbach, M., Eitemüller, C., Rodemer, M., & Walpuski, M. (2025). The influence of representations on task difficulty in organic chemistry: an exploration using a novel paired-items test instrument. *International Journal of Science Education*, *47*(14), 1941–1963. https://doi.org/10.1080/09500693.2024.2378218

Szulewski, A., Braund, H., Egan, R., Hall, A. K., Dagnone, J. D., Gegenfurtner, A., & van Merrienboer, J. J. G. (2018). Through the Learner's Lens: Eye-Tracking Augmented Debriefing in Medical Simulation. *Journal of Graduate Medical Education*, *10*(3), 340–341. https://doi.org/10.4300/jgme-d-17-00827.1

Taconis, R., Ferguson-Hessler, M. G. M., & Broekkamp, H. (2001). Teaching science problem solving: An overview of experimental work. *Journal of Research in Science Teaching*, *38*(4), 442–468. https://doi.org/10.1002/tea.1013

Tang, H., Kirk, J., & Pienta, N. J. (2014). Investigating the effect of complexity factors in stoichiometry problems using logistic regression and eye tracking. *Journal of Chemical Education*, *91*(7), 969–975. https://doi.org/10.1021/ed4004113

Tang, H., & Pienta, N. (2012). Eye-tracking study of complexity in gas law problems. *Journal of Chemical Education*, *89*(8), 988–994. https://doi.org/10.1021/ed200644k

Theeuwes, J. (2019). Goal-driven, stimulus-driven, and history-driven selection. *Current Opinion in Psychology*, *29*, 97–101. https://doi.org/10.1016/j.copsyc.2018.12.024

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20*(2), 147–168. https://doi.org/10.1080/13546783.2013.844729

Tóthová, M., Rusek, M., & Chytrý, V. (2021). Students' procedure when solving problem tasks based on the periodic table: An eye-tracking study. *Journal of Chemical Education*, *98*(6), 1831–1840. https://doi.org/10.1021/acs.jchemed.1c00167

Tsai, P.-Y., Yang, T.-T., She, H.-C., & Chen, S.-C. (2019). Leveraging College Students' Scientific Evidence-Based Reasoning Performance with Eye-Tracking-Supported Metacognition. *Journal of Science Education and Technology*, *28*(6), 613–627. https://doi.org/10.1007/s10956-019-09791-x

van Gog, T., Janssen, E., Lucas, F., & Taheij, M. (2024). A Motivational Perspective on (Anticipated) Mental Effort Investment: The Biopsychosocial Model of Challenge and Threat. *Educational Psychology Review*, *36*(2), 54. https://doi.org/10.1007/s10648-024-09861-6

van Gog, T., Paas, F., Van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

van Gog, T., & Rummel, N. (2010). Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review*, *22*(2), 155–174. https://doi.org/10.1007/s10648-010-9134-7

Wang, J.-R., Wang, Y.-C., Tai, H.-J., & Chen, W.-J. (2010). Investigating the Effectiveness of Inquiry-Based Instruction on Students With Different Prior Knowledge and Reading Abilities. *International Journal of Science and Mathematics Education*, *8*(5), 801–820. https://doi.org/10.1007/s10763-009-9186-7

Ward, L., Rotich, F., Raker, J. R., Komperda, R., Nedungadi, S., & Popova, M. (2025). Design, development, and evaluation of the organic chemistry representational competence assessment (ORCA). *Chemistry Education Research and Practice*, *26*(1), 244–258. https://doi.org/10.1039/D3RP00188A

Weinrich, M., & Britt, R. (2022). Students' Attention on Curved Arrows While Evaluating the Plausibility of an Organic Mechanistic Step. In N. Graulich & G. Shultz (Eds.), *Student Reasoning in Organic Chemistry* (pp. 3–18). The Royal Society of Chemistry. https://doi.org/10.1039/9781839167782-00001

Weinrich, M. L., & Sevian, H. (2017). Capturing students' abstraction while solving organic reaction mechanism problems across a semester. *Chemistry Education Research and Practice*, *18*(1), 169–190. https://doi.org/10.1039/C6RP00120C

Weir, J. J. (1974). Problem Solving Is Everybody's Problem. The Science Teacher, 41(4), 16–18.

Wernecke, U., Schütte, K., Schwanewedel, J., & Harms, U. (2018). Enhancing Conceptual Knowledge of Energy in Biology with Incorrect Representations. *CBE—Life Sciences Education*, *17*(1), ar5. https://doi.org/10.1187/cbe.17-07-0133

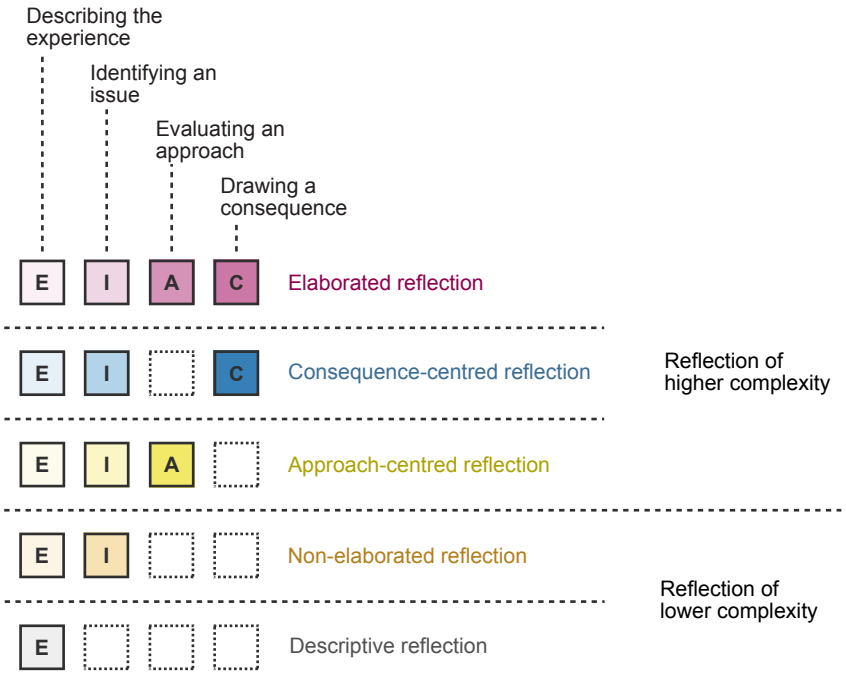
Williamson, V. M., Hegarty, M., Deslongchamps, G., Williamson III, K. C., & Shultz, M. J. (2013). Identifying student use of ball-and-stick images versus electrostatic potential map images via eye tracking. *Journal of Chemical Education*, *90*(2), 159–164. https://doi.org/10.1021/ed200259j

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2023). Utilizing a Pretrained Language Model (BERT) to Classify Preservice Physics Teachers' Written Reflections. *International Journal of Artificial Intelligence in Education*, *33*(3), 439–466. https://doi.org/10.1007/s40593-022-00290-6

Ye, J., Zheng, Y., Zhan, M., Zhou, Y., Li, L., & Chen, D. (2024). Characteristics of Pre-Service Chemistry Teachers' Mechanistic Reasoning In Organic Chemistry Tasks: An Eye-Tracking Study. *Research in Science Education*, *55*, 277–296. https://doi.org/10.1007/s11165-024-10185-2

Yuriev, E., Naidu, S., Schembri, L. S., & Short, J. L. (2017). Scaffolding the development of problem-solving skills in chemistry: guiding novice students out of dead ends and false starts. *Chemistry Education Research and Practice*, *18*(3), 486–504. https://doi.org/10.1039/C7RP00009J

STUDENTS' REFLECTION



Langner, A., & Graulich, N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
https://doi.org/10.1080/09500693.2024.2430804

From sight to insight – reflection processes in an eye-gaze-augmented retrospective

Axel Langner and Nicole Graulich

Institute of Chemistry Education, Justus-Liebig-University Giessen, Giessen, Germany

ABSTRACT

Various one-size-fits-all methods have been developed to improve students' use of representations for problem-solving. However, as the interplay of processing representations and problem-solving might be highly individual, a differentiated approach may be needed to foster students' problem-solving with representations. A retrospective enhanced with eye-gaze replays and reflection prompts (eye-gaze-augmented retrospective) may stimulate students' reflection on their problem-solving process by providing normally inaccessible information, allowing the students to look back and draw tailored consequences for future problem-solving situations. Therefore, an exploratory study was conducted where students solved chemistry problems and subsequently reflected on their problem-solving while watching their eye-gaze replays in three distinct reflection phases. The study aimed to investigate to what extent different phases of an eye-gaze-augmented retrospective stimulate students' reflection on their problem-solving by characterizing the structural complexity of the reflection processes and considering the accuracy of students' problem-solving. The results showed that reflection processes with varying structural complexity emerged. Higher complexity reflections emerged primarily in the phase in which the task solution was presented accompanied by reflection prompts, especially for inaccurately solved problems. The findings suggest that an eye-gaze-augmented retrospective could stimulate reflection on problem-solving with representations, but optimization is necessary to enhance each student's reflection process effectively.

ARTICLE HISTORY
Received 2 May 2024
Accepted 13 November 2024

KEYWORDS
Eye-tracking; reflection; organic chemistry

1. Introduction

In science disciplines, problem-solving is often linked to representational competence (Kozma & Russell, 2005; Prain & Tytler, 2012; Schönborn & Anderson, 2009). Consequently, students' use of representations is integral to developing their expertise in solving problems. However, several studies indicated that students often experience

CONTACT Nicole Graulich Nicole.Graulich@didaktik.chemie.uni-giessen.de Institute of Chemistry Education, Justus-Liebig-University Giessen, Heinrich-Buff-Ring 17, Giessen 35392, Germany

challenges in solving problems with representations in different science disciplines, leading them to apply rules or heuristics, and to struggle to derive implicit information, often due to inadequately integrating different conceptual information or insufficient prior knowledge (e.g. Bollen et al., 2016; Bollen et al., 2017; Graulich et al., 2019; Halverson et al., 2011; McClary & Talanquer, 2011; Susac et al., 2023). Additionally, several eye-tracking studies in science education have shown that students who are unsuccessful or have low prior knowledge often focus on irrelevant aspects or show unproductive visual behavior (e.g. Braun et al., 2022; Klein et al., 2018; Rodemer et al., 2020; Susac et al., 2023; Tóthová & Rusek, 2022). For example, in a study by Braun et al. (2022), the structural features students attended to and their visual selection approach while drawing resonance structures were investigated. While looking at an organic representation, unsuccessful students distributed their attention more to irrelevant structural features and focused on single structural features in their reasoning, rather than on relevant interrelated features necessary to solve the task successfully. Furthermore, the results of the study indicated that although successful and unsuccessful students showed differing viewing behavior, both types of students required additional time to decode and identify relevant structural features when the organic representations became more complex. In addition, in an instruction-based study by Klein et al. (2018), students were introduced to two visual strategies to qualitatively interpret the divergence of graphical vector fields. Students' performance was the lowest when they were introduced to only one of the strategies, rather than both, and were unable to choose their strategy freely. In a comparison between the most unsuccessful and most successful students, the results demonstrated that the most unsuccessful students exhibited visual behavior misaligned with respect to the representation. Considering the importance of problem-solving and the challenges students experience, one main goal of instruction in science education is to support students in successfully using representations for problem-solving.

Various, often one-size-fits-all, methods to improve students' use of representations, visually and/or conceptually, are well-documented, such as cueing, example-based learning, or scaffolding (e.g. Alfieri et al., 2013; Alpizar et al., 2020; Emhardt et al., 2023; Frey et al., 2022; van Gog & Rummel, 2010). For instances, Jarodzka et al. (2013) demonstrated that eye-movement-modelling examples can guide students' visual attention to successfully classify fish locomotion patterns, Rodemer et al. (2021) and Rodemer et al. (2022) reported that highlighting helped students to better focus their attention on the relevant features of the representations while also decreasing extraneous cognitive load and compensating low prior knowledge, and Kranz et al. (2023) showed that using scaffolded contrasting cases of two organic reaction mechanisms improved the knowledge integration of students with low prior knowledge. However, as studies, for example by Braun et al. (2022) and Klein et al. (2018), demonstrated, the interplay of processing representations and one's problem-solving process is highly complex and individual. Students vary in many aspects, such as prior knowledge or problem-solving strategies, which play a crucial role in processing representations (Baluyut & Holme, 2019; Cook et al., 2008). Thus, to support students in their visual and conceptual processing during problem-solving, while considering their unique experiences, prior knowledge, strategies, and challenges with decoding representations, more individualized and differentiated support is needed. Encouraging students to reflect on their problem-solving approaches may be such individual support, as reflection is an

important key for more personalized learning (Shemshack et al., 2021). So far, little attention has been devoted to more differentiated instructional approaches that allow students to reflect on their problem-solving approaches – especially focusing on the visual *and* conceptual processing during their problem-solving. This process may be enhanced *conceptually* by a combination of reflection prompts or the task solution and *visually* by one's own eye movements that usually would not be accessible to a student (Fleck & Fitzpatrick, 2010).

1.1. Reflection for learning

While performing an action (i.e. solving a problem), the knowledge applied is generally implicit in both the patterns of action and emotions while doing it (Schön, 1983). Schön (1983) described the introspective process of examining this action and the knowledge it implies as *reflection*. The intended outcome of such a process is to guide future actions (Killion & Todnem, 1991).

Several hierarchical models were developed to characterize reflection processes (e.g. Nguyen et al., 2014). They have in common that they differentiate between rather descriptive or highly elaborated ways of reflection and different characteristics of reflection to characterize reflection processes. These models address reflection either in general or in specific contexts such as teacher education or medical education. Some of the therein discussed elements of reflection in these models can be adapted to the context of problem-solving. For example, the aspect of moving back to a past experience (e.g. by describing the problem-solving process) (see Korthagen & Kessels, 1999), experiencing an issue, difficulty, or challenge (e.g. by identifying issues that occurred while problem-solving) (see Boyd & Fales, 1983; Moon, 2004), and taking other points of view or methods of actions into account (e.g. by evaluating other approaches to solve the problem) (see Fleck & Fitzpatrick, 2010; Korthagen, 1985). Furthermore, the final element of reflection is to draw consequences, as reflection aims to lead to new perspectives, improvements, or a greater understanding of an experience (see Hatton & Smith, 1995; Moon, 2004; Sandars, 2009). In problem-solving, this might be related to recalibrating one's own approach by formulating new approaches or thoughts that are important to solve a future problem.

Various models identified the role of reflection in learning and emphasized the importance of reflective processes for learning (e.g. Boud et al., 1985; Kolb & Fry, 1975), and other studies indicated that more elaborated reflection processes lead to more effective learning (e.g. Corrales & Erwin, 2020). The potential of reflection can be further supported by an assisting educator following a reflection-supporting model or by using tools that allow insights into one's behavior or problem-solving, i.e. sensor devices (Boud et al., 1985; Fleck & Fitzpatrick, 2010).

1.2. Supporting reflection with eye-tracking technology

Different reflection-supporting models were developed to facilitate reflection in various professions with specific foci, e.g. the ALACT model of Korthagen (1985) for teacher education. Even though the ALACT model was originally developed for reflection on teaching situations, this model is also suited for problem-solving situations, as it

focuses on essential aspects of an action and structures the reflection process to develop new courses of action for future trials. The ALACT model distinguishes three phases of reflection in a reflection-based learning scenario: *looking back on the action*, *awareness of essential aspects*, and *creating alternative methods of action*, which are initiated after an *action* has been completed and before embarking on a new *trial*.

To support students' reflection, i.e. to look back on the action, be aware of essential aspects, and create alternative methods of action, purposeful reflection prompts can guide such reflective activities (Guo, 2022). Especially to support the act of looking back and being aware of essential aspects, sensor devices could be used to represent new information that otherwise would not be available and to offer an additional perspective (Fleck & Fitzpatrick, 2010). One of these sensor devices can be an eye-tracker that records where, for how long, and in which order an individual looks at a visual representation. An eye-gaze replay, depicting conscious and unconscious eye movements dynamically, could be used to offer students new information about their visual and conceptual processing of a complex representation, i.e. their visual problem-solving actions (see Hyrskykari et al., 2008; van Gog et al., 2005). In this context, eye-tracking is not used as an assessment tool to evaluate students' learning based on attention distribution on representations. Instead, it is used as a feedback tool to enhance students' reflection processes and their subsequent learning. So far, only a few studies have used eye-tracking as a feedback tool in the context of problem-solving in science education. In a study by Hansen et al. (2019), for example, students had to critique chemically relevant features of accurate and inaccurate animations and static images of redox and precipitation reactions. By presenting students with static inverse opacity images of their eye movements, Hansen et al. (2019) demonstrated that this type of eye-tracking feedback can be used to make students aware of their viewing behavior. Additionally, Tsai et al. (2019) showed that eye-gaze replays can be used to foster metacognition on scientific evidence-based reasoning tasks to improve students' performance.

Hence, a retrospective cued by eye-gaze replays and accompanied by reflection prompts (referred to as an eye-gaze-augmented retrospective; for detailed information, see section 2.3.) could allow students to revisit their problem-solving approach in detail, provide students with additional individual information, make them aware of crucial aspects of their unique process, and allow them to draw tailored consequences to inform future trials. However, to what extent eye-gaze-augmented retrospectives stimulate reflection processes in the context of problem-solving with chemical representations has not yet been investigated.

1.3. Purpose of this study

This qualitative, exploratory study aims to investigate the extent to which an eye-gaze-augmented retrospective can stimulate students' reflection on their problem-solving with representations. Given the complex and multifaceted nature of representations in organic chemistry, this domain provides a suitable context for problem-solving tasks and might allow implications for other science disciplines working with complex representations. Therefore, we characterized the reflection processes of undergraduate organic chemistry students on typical structural sites within representations of organic chemistry. Specifically, we were interested in the structural complexity of the reflection

process, i.e. the co-occurrence of reflection elements. As the eye-gaze-augmented retrospective was structured to be aligned with the three reflection phases outlined in the ALACT model (*looking back on the action*, *awareness of essential aspects*, and *creating alternative methods of action*), we were further interested in the contribution of these phases and how they might stimulate more (or less) structurally complex reflection processes. Therefore, the following research questions guided our investigation:

(1) How is the structural complexity of reflection processes in organic chemistry characterized when students' reflection processes are stimulated by an eye-gaze-augmented retrospective?
(2) To what extent are reflection patterns of higher complexity stimulated by certain phases along the eye-gaze-augmented retrospective and are dependent on student performance?

2. Material and methods

2.1. Context and participants

Data were collected in 2021 at a German university. At that time, the organic chemistry 1 course enrolled around 60 students. After visiting the course, twenty-two students voluntarily participated in this study. In the recruitment process, no pre-selection was made, and recruitment continued until data saturation was reached, i.e. further data did not show new observations across the phases of the intervention. The lecture in this course followed a traditional curriculum and was structured by functional groups, mechanism types, and structure-property-relationships. The topics relevant to this study (R/S configuration of asymmetric carbon atoms, Brønsted acid-base concept, and electrophilicity and nucleophilicity) were covered by the lecture. Sixteen participants identified as female, and six participants as male. For this publication, participants were named with pseudonyms according to their self-reported gender identity. The average age of the participants was 24.0 ± 3.4 years. The participants were native German speakers who had normal or corrected-to-normal vision. The data collection was conducted in German. Prompts and quotes were translated carefully for this publication to ensure that the wording and meaning were represented appropriately.

2.2. Study design

2.2.1. Procedure

Data was collected in a single eye-tracking and interview session as part of a larger study. Each participant underwent the following procedure for each of the three tasks: The participants were first prompted to mentally solve the presented task (Figure 1, bottom). While solving a task, the eye movements of the participants were tracked. Immediately after the participants solved the task, they noted their responses on a sheet of paper on which the respective molecule was depicted. Since the eye-gaze-augmented retrospective included a retrospective think-aloud, the participants did not need to explain their responses at this stage. After each task, the phases of the eye-gaze-augmented retrospective were carried out (Figure 1, top). During the eye-gaze-augmented retrospective,

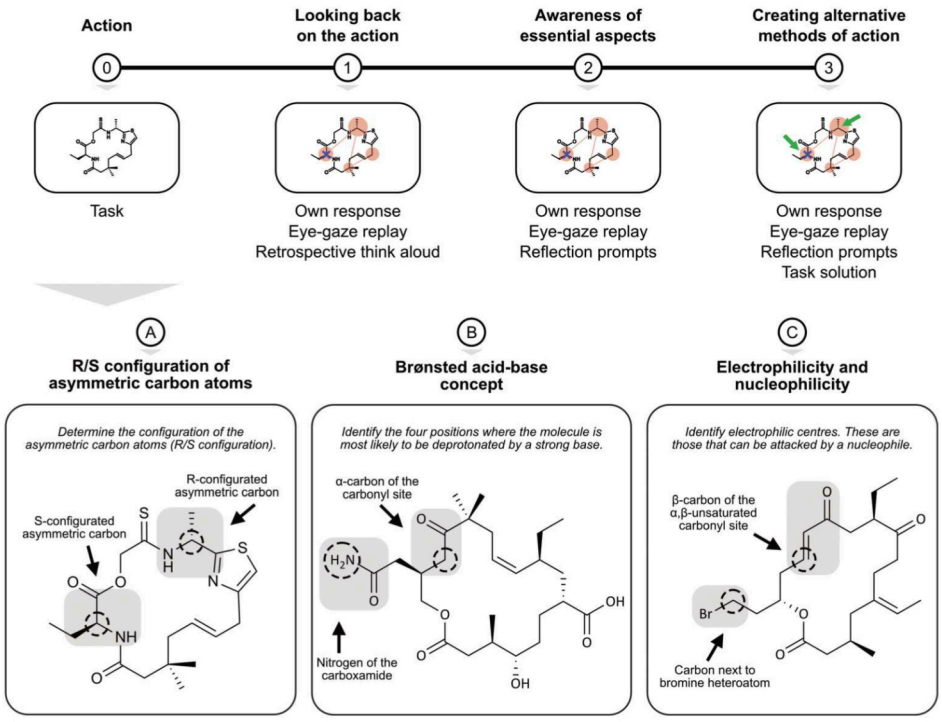


Figure 1. Study design with the phases of the retrospective (upper half) and the three tasks (lower half). Two typical structural sites of each task context were chosen for data analysis (highlighted in grey).

the participants were audiotaped, and their interaction with their eye-gaze replay was videotaped. By doing this, participants reflected on their problem-solving process on different structural sites of the molecules. For the subsequent analysis, two typical structural sites relevant to the specific problem-solving task at hand were selected (for example, in task C, concerning electrophilicity and nucleophilicity, the carbon next to the halogen heteroatom would represent a typical structural site that needs to be identified) (Figure 1, bottom).

Prior to the main study, a pilot study with a small sample was conducted to ensure the clarity of the tasks and prompts, and to test and refine the data collection procedures. Insights from the pilot were used to adjust the final study design, ensuring alignment with the research objectives. Additionally, the study design was discussed with experts in this discipline.

2.2.2. Task design

The tasks in this study were chosen to be representative of the complexity of a lecture for organic chemistry beginners. For this purpose, task prompts were chosen from three common topics: *Determine the configuration of the asymmetric carbon atoms (R/S configuration)* for R/S configuration of asymmetric carbon atoms (Figure 1, task A), *Identify the four positions where the molecule is most likely to be deprotonated by a*

strong base for the Brønsted acid–base concept (Figure 1, task B), and *Identify electrophilic centres. These are those that can be attacked by nucleophiles* for electrophilicity and nucleophilicity (Figure 1, task C).

To ensure that the task difficulty was appropriate for the prior knowledge of the students enrolled in the organic chemistry 1 course, we discussed the appropriateness of the task design with the course instructors beforehand.

2.2.3. Eye-gaze replay

Eye movements of the participants were tracked in the *action* phase (Figure 1, top) using the software Tobii Pro Lab, a Tobii Pro X3-120 screen-based eye-tracking system (120 Hz) connected to a Tobii Pro External Processing Unit in a live viewer setup with a 24-inch presentation monitor (1920 × 1200). To reduce data loss, the participants were briefed to minimize head and body movements during the recording and to remain seated at a distance of 60–65 cm from the presentation screen. Before each task, the eye-tracking system was calibrated with nine points and validated with four points. Then, the task prompt was displayed first. By pressing a key, a fixation cross appeared for two seconds before the corresponding molecule was shown (Figure 1, bottom). The molecules of each task were presented in a size that the distances between atoms were approximately 2.2 cm. This corresponds to an eye movement of approximately 2°, exceeding the accuracy of the eye tracker of 0.7°. With this, participants should be able to allocate their fixation points in the eye-gaze replay to viewed atoms.

Immediately after the eye-tracking data collection, the eye-gaze replay was exported from Tobii Pro Lab, imported into the video-editing software Hitfilm Express, and overlaid with the student's response for the first and second phases of the retrospective. For the third phase, the task solution was additionally added as an overlay to the eye-gaze replay and the participants' task solution.

2.3. Eye-gaze-augmented retrospective

2.3.1. Design and key considerations

The eye-gaze-augmented retrospective was structured as a semi-structured interview, which allowed the interviewer to specifically address certain individual events described by the participants. The interviewer used a mix of generic prompts to encourage thinking aloud and direct prompts for specific, contextualized questions (Davis, 2003). In doing so, the interviewer included both low-level questions for memory recall and high-level questions for comprehension and connection (King, 1994). In the development of the reflection prompts for the eye-gaze-augmented retrospective, we followed suggestions from Ericsson and Simon (1993) and Korthagen and Kessels (1999), like, asking for alternatives and requesting consequences for the next time. The given prompts aimed to trigger and guide the reflection process (Moon, 2004). Additionally, we ensured that the reflection prompts were non-judgmental to provide a supportive environment (Sandars, 2009). Furthermore, suggestions for technology-enhanced reflection were considered for the design and conception of the retrospective, like asking for explanations and giving the opportunity to re-watch the replay (Fleck & Fitzpatrick, 2010).

Additionally, to guide the reflection in the eye-gaze-augmented retrospective, we followed the reflection phases of the ALACT model of Korthagen (1985) (Figure 1, top).

2.3.2. First phase (looking back on the action)

In the first phase, the participants watched the eye-gaze replay of their problem-solving superimposed with their given response and described their problem-solving in a retrospective think-aloud. The replay speed was reduced to half the normal speed as eye movements are very fast, and participants should be able to express their thoughts regarding their actions (Hyrskykari et al., 2008). The playback times of the replays with reduced speed were 4:15 ± 2:15 min. The participants were instructed that they could pause the eye-gaze replay, skip forward or backward, and rewatch the replay as many times as they wanted. Furthermore, the elements of the eye-gaze replay (i.e. fixations and saccades) were explained. Participants were prompted to describe their experience as detailed as possible with the help of the eye-gaze replay. The instruction here followed the standards for retrospective think-aloud protocols, like prompting to report everything, requesting to express memory uncertainties, and asking not to work on the problem again (Ericsson & Simon, 1993). If participants omitted or just briefly described events, the interviewer asked for a detailed description.

2.3.3. Second phase (awareness of essential aspects)

In the second phase, the participants rewatched the eye-gaze replay and were additionally guided by reflection prompts to identify issues, evaluate approaches, or draw consequences (e.g. *Is there anything you notice?*, *What is the reason for … ?*, or *Would you do anything conceptually and visually different if you had another chance to solve this or a similar problem again?).*

2.3.4. Third phase (creating alternative methods of action)

In the third phase, the task solution was additionally superimposed in the replay. Again, the participants were guided by similar reflection prompts (e.g. *What do you conclude from the task solution? What is the reason for … ?* or *Would you now change anything conceptually and visually if you had another chance to solve this or a similar problem again?).* This allowed participants to compare their eye movements and responses with the task solution to identify productive and unproductive behavior. Only in the third phase did the participants see the task solution, but no further conceptual explanation, e.g. why these structural sites are accurate, was given at any time.

2.4. Data analysis

2.4.1. Data preparation

As we focus on reflection processes on two typical structural sites for each task (Figure 1, bottom), our analysis resulted in a total sum of 132 single reflection processes. The interview data of the retrospective were transcribed verbatim, and text passages in which participants addressed those structural sites were selected for qualitative content analysis. On average, participants reflected 4:42 ± 1:30 min on their problem-solving at one structural site. Additionally, participants' responses at each structural site were categoried as either accurate or inaccurate.

2.4.2. Characterizing the structural complexity of reflection processes

To characterize the structural complexity of the reflection processes on problem-solving (research question 1), we coded the reflection elements in the interview data via deductive qualitative content analysis and grouped them in patterns based on the co-occurrence of the reflection elements. In doing so, the coding was informed by the broad theory of reflection (Boyd & Fales, 1983; Fleck & Fitzpatrick, 2010; Hatton & Smith, 1995; Korthagen, 1985; Korthagen & Kessels, 1999; Moon, 2004; Nguyen et al., 2014). In total, four codes were used: *Describing the experience* (E), *identifying an issue* (I), *evaluating an approach* (A), and *drawing a consequence* (C) (Table 1). This approach allowed for the systematic identification and coding of specific reflection elements grounded in literature.

As the analysis focused on the structural complexity, we coded each reflection process dichotomously, noting the presence or absence of a code.

The coding scheme was constantly discussed between the authors to ensure that coding decisions represented the data. For interrater reliability, an undergraduate research assistant coded a random sample of 20% of the data independently. A kappa coefficient κ_n of 0.95 was reached (Brennan & Prediger, 1981), indicating high agreement and reliability for the coding scheme (Kuckartz & Rädiker, 2019).

2.4.3. Development of the structural complexity of reflection patterns along the retrospective

To analyse which phases of the eye-gaze-augmented retrospective might stimulate reflection patterns of higher complexity (research question 2), we quantified the change of reflection patterns across the three phases. In this process, the codes of the reflection elements from each phase were not considered in isolation. Instead, they were cumulatively added to the existing codes from the previous phases, allowing us to characterize the development of the reflection process throughout the retrospective while considering its dynamic nature. This cumulative approach was chosen because participants did not necessarily repeat their mentioned experiences, issues, approaches, or consequences in each phase if they had already been mentioned in previous phases, enabling us to capture the ongoing development of the reflection processes without losing information from earlier phases. By comparing which reflection elements were added in each phase of the retrospective, we could infer which phases may stimulate the occurrence of specific reflection elements. Furthermore, as a reflection process can depend on contextual

Table 1. Description of the codes, representing reflection elements, used to characterize the structural complexity of the reflection processes.

Code	Code description
Describing the experience (E)	The participant describes or recalls the problem-solving process (e.g. related to chemical concept knowledge, visual behavior, or general thoughts).
Identifying an issue (I)	The participant verbalizes difficulties, challenges, or uncertainties they had or still have (e.g. related to chemical concept knowledge, visual behavior, or general behavior).
Evaluating an approach (A)	The participant reviews their assumptions and behavior, or tests alternative assumptions and behavior (e.g. related to chemical concept knowledge, visual behavior, or general behavior).
Drawing a consequence (C)	The participant verbalizes intentions or changes regarding their actions and thoughts with reference to future problem-solving situations (e.g. related to chemical concept knowledge, visual behavior, general behavior, or general thoughts).

factors (Boud et al., 1985), the accuracy of determining the structural site was additionally considered.

3. Results & discussion

3.1. How is the structural complexity of reflection processes in organic chemistry characterized when students' reflection processes are stimulated by an eye-gaze-augmented retrospective?

3.1.1. Distinct reflection patterns

Using the four codes representing reflection elements proposed in Table 1 led to the emergence of five distinct patterns of the reflection processes that occurred during the eye-gaze-augmented retrospective with varying degrees in the structural complexity from descriptive to elaborated reflection pattern (Figure 2).

Depending on the type of reflection elements present, these reflection patterns can be characterized from less to more complex reflection patterns in five categories: *descriptive reflection*, *non-elaborated reflection*, *approach-centred reflection*, *consequence-centred reflection,* and *elaborated reflection* (Table 2).

3.1.2. Reflection processes of lower complexity

We frame the *descriptive reflection* pattern as the lowest level of reflection processes (e.g. Steven's example, Figure 3), although the literature is unclear if moving back to experiences or turning one's thoughts back on action counts as an explicit part of the reflection process (Killion & Todnem, 1991; Schön, 1983). Others argue that a superficial

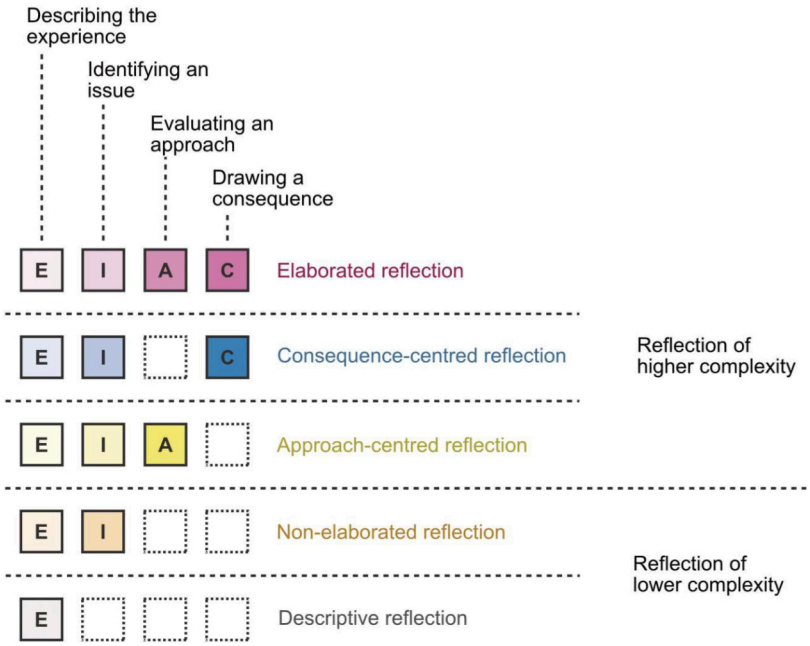


Figure 2. Emerged reflection patterns with varying structural complexity.

Table 2. Characterization of the reflection patterns stimulated by the eye-gaze-augmented retrospective.

Reflection pattern	General description
Elaborated reflection	The reflection process involves describing experiences and issues, evaluating one's own or an alternative approach, and verbalizing intentions or changes with reference to future problem-solving. (Code E, I, A, C)
Consequence-centred reflection	The reflection process involves describing experiences and issues, and additionally verbalizes intentions or changes with reference to future problem-solving. (Code E, I, C)
Approach-centred reflection	The reflection process involves describing experiences and issues, and additionally evaluates problem-solving by reviewing or testing an alternative approach. (Code E, I, A)
Non-elaborated reflection	The reflection process involves describing experiences and additionally expresses issues encountered during problem-solving or that still persist. (Code E, I)
Descriptive reflection	The reflection process involves describing experiences. (Code E)

description without any further elaboration, like in the *descriptive reflection* pattern, cannot be considered a reflection (Fleck & Fitzpatrick, 2010; Hatton & Smith, 1995). Nevertheless, different reflection models identified looking back on thoughts and actions as a necessary part of the reflection process (e.g. Korthagen & Kessels, 1999; Nguyen et al., 2014). As all students were prompted to describe their thoughts and actions during the retrospective, all started with describing their problem-solving process and, thus, reached the first reflection pattern.

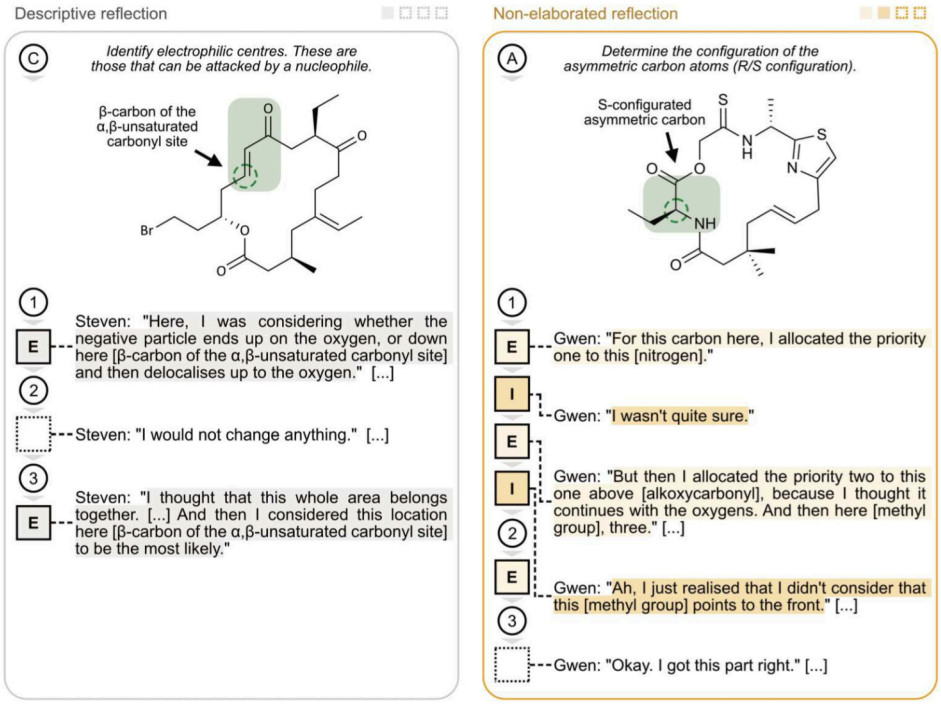


Figure 3. Excerpts of a descriptive reflection pattern at the β-carbon of the α,β-unsaturated carbonyl site of the task about electrophilicity and nucleophilicity (left, Steven's example) and a non-elaborated reflection pattern at the S-configurated asymmetric carbon of the task about R/S configuration of asymmetric carbon atoms (right, Gwen's example). Both determined the structural sites accurately.

Similar to the *descriptive reflection* pattern, the *non-elaborated reflection* pattern mainly contained a description of experiences but also identified difficulties and challenges students experienced or still face. Nevertheless, students who showed the *non-elaborated reflection* pattern did not further expand their reflection on other reflection elements, i.e. evaluating approaches or drawing consequences (e.g. Gwen's example, Figure 3). Therefore, the patterns *descriptive reflection* and *non-elaborated reflection* can be considered as reflection processes of lower complexity (Figure 2).

Figure 3 illustrates two students' examples of lower reflection complexity. Steven's reflection process was categorized as a *descriptive reflection* pattern at this respective structural site (i.e. the β-carbon of the α,β-unsaturated carbonyl site in task C), whereas Gwen showed a *non-elaborated reflection*. In his reflection, Steven explicitly stated that he would not change anything regarding his problem-solving strategy in the second phase as he was certain about his problem-solving strategy and, in the end, determined the structural site accurately. He seemed pleased with his problem-solving strategy and result and saw no need to evaluate different approaches or draw consequences.

In comparison to Steven's reflection, which only describes his problem-solving process, Gwen additionally identified issues within her problem-solving process during her reflection. Gwen identified two issues in her problem-solving approach: first, she was not sure about the allocation of priorities, and second, she did not consider the spatial arrangement of the carbon's residues. Nevertheless, Gwen's reflection process ended after she had verified that her response was accurate. Although her issues did not influence the accuracy of her response in this case, she did not further explore the reasons for this issue. Gwen might have benefited more in this retrospective by evaluating different approaches and drawing consequences to be prepared for future problem-solving situations. However, she was pleased that her response was accurate and saw no need to reflect further. In her case, the retrospective did not stimulate an advanced reflection process, although an unexpected success, like in her case, could be a stimulus for further reflection (Schön, 1983).

Both Steven and Gwen appeared to see no need for further reflection in their reflection processes. Similarly, in a study by Blackford et al. (2023), organic chemistry students reported that one reason for not reflecting on the outcome or process was that they did not consider it necessary once they had already found a solution. This perspective might be reinforced when students' responses are confirmed by the task solution, diminishing the perceived need for deeper reflection.

These two cases are examples of reflection patterns that did not further include the evaluation of different approaches or drawing consequences.

3.1.3. Reflection processes of higher complexity

The reflection patterns *approach-centred reflection, consequence-centred reflection,* and *elaborated reflection* went beyond describing the experience or identifying issues. Students showing these reflection patterns were further engaged in evaluating their own or proposing alternative approaches, in formulating consequences, or in doing both. As these reflection patterns included multiple reflection elements, we considered them reflection processes of higher complexity (Figure 2). The reflection element *identifying an issue,* also present in reflection processes of lower complexity, appears in every

reflection process of higher complexity, showing that identifying an issue seems to be a potential prerequisite for reflection processes of higher complexity. This aligns with observations made by others suggesting that uncertainty (Dewey, 1933; Schön, 1983), as well as related experiences such as doubt, confusion, concerns of not knowing, and lack of information in explaining something (for a review, see Marshall et al., 2021), can act as a trigger for reflection. In their reflection processes, students in this study reported issues consistent with difficulties documented in the science education literature, such as challenges in deriving implicit information, insufficient prior knowledge, or unproductive visual behavior (e.g. Bollen et al., 2015; Cooper et al., 2010; Graulich et al., 2019; Klein et al., 2018). Building upon their identified issues, in reflection processes of higher complexity, students could go deeper in their reflection. For example, by reviewing their assumptions or testing alternative ones, like Peter in his *approach-centred reflection* (Peter's example, Figure 4), or by formulating consequences, like Susan in her *consequence-centred reflection* (Susan's example, Figure 4).

A longer fixation on the double bond in his eye-gaze replay initiated Peter's reflection process (Peter's example, Figure 4). After seeing the task solution in the third phase of the retrospective, he started with a conceptual sense-making process to explain why he failed to identify the structural site accurately. He elaborated on alternative ways of looking at this structural site and expressed how bonds could flip. However, he could not finally

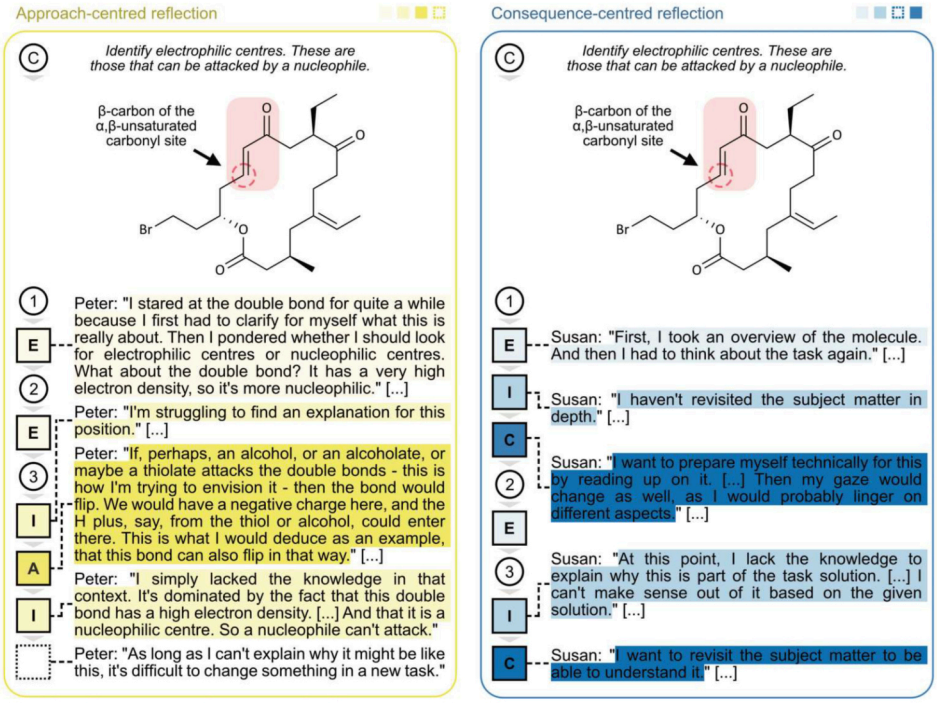


Figure 4. Excerpts of an approach-centred reflection pattern (left, Peter's example) and a consequencecentred reflection pattern (right, Susan's example) at the β-carbon of the α,β-unsaturated carbonyl site of the task about electrophilicity and nucleophilicity. Both did not determine this structural site accurately.

deduce if this double bond is an electrophilic or a nucleophilic centre, as he did not consider the influence of the carbonyl group. He clearly expressed that conceptual knowledge is missing, which did not allow him to formulate a consequence. Peter's *approach-centred reflection* is an example of students being aware of their lack of knowledge, which limits them to draw consequences. Other students showing reflection processes of higher complexity did not as explicitly state this as Peter did. Instead, they stopped their reflection process or drew consequences that were not directly connected to an immediate change in their knowledge, behavior, or thoughts. For example, based on her lack of chemical concept knowledge, Susan, in her *consequence-centred reflection* (Susan's example, Figure 4), drew the consequence of 'learning the content' already in the first phase of the retrospective. The visual information of the task solution in the third phase alone was not sufficient for her to find an explanation for it, and she remained with her initial consequence. Both examples show that for some students, further conceptual support is necessary to help students test assumptions and formulate consequences purposefully.

Anna, as an example of an *elaborated reflection* (Anna's example, Figure 5), seems to have had sufficient chemical concept knowledge to be able to make sense of the presented task solution, leading to purposeful consequences. In the third phase of the retrospective, Anna recognized that she had mistaken an amine for an amide and that, even as an amide, she would not have determined it as an electrophilic centre. After considering an alternative visual behavior (e.g. *I should have looked more closely*) and testing a conceptual assumption (e.g. *But thinking about it now, if I have an amide, so there is a carbonyl group* […]), she drew the consequence of examining more adjacent atoms and not to focus only on hydrogens anymore. Overall, the extent to which a consequence could be drawn, and consequently, the reflection pattern that emerges, seems to depend partially on students' ability to activate the conceptual knowledge to make sense of their problem-solving and the task solution.

Anna's example is also exemplary for students who formulated more visual-related consequences in their reflection, stimulated by the retrospective (e.g. focusing on structural features or looking closely or stepwise), even though conceptual issues were identified. The retrospective provides only visual-related information, i.e. the eye-gaze replay and the visual depiction of the task solution. This framing might stimulate students to focus more on their visual than their conceptual processing. A similar observation can be seen in Felicia's *elaborated reflection* (Felicia's example, Figure 5). In contrast to Anna's reflection, Felicia accurately determined the structural site. In the second phase of the retrospective, she expressed dissatisfaction with her problem-solving approach and drew a general consequence for her future behavior. When asked again if she would draw a consequence in the third phase, she added the visual consequence to look more successively in future problem-solving situations. This highlights that students were stimulated to focus more on their visual behavior due to the visual framing of the retrospective.

The goal of the reflection in this retrospective is to guide future problem-solving situations by drawing purposeful consequences. Although the identification of issues was a precursor for the evaluation of approaches and the drawing of consequences, with the pattern of *consequence-centred reflection*, the results showed that it is possible to draw consequences without necessarily evaluating different approaches. In the *approach-*

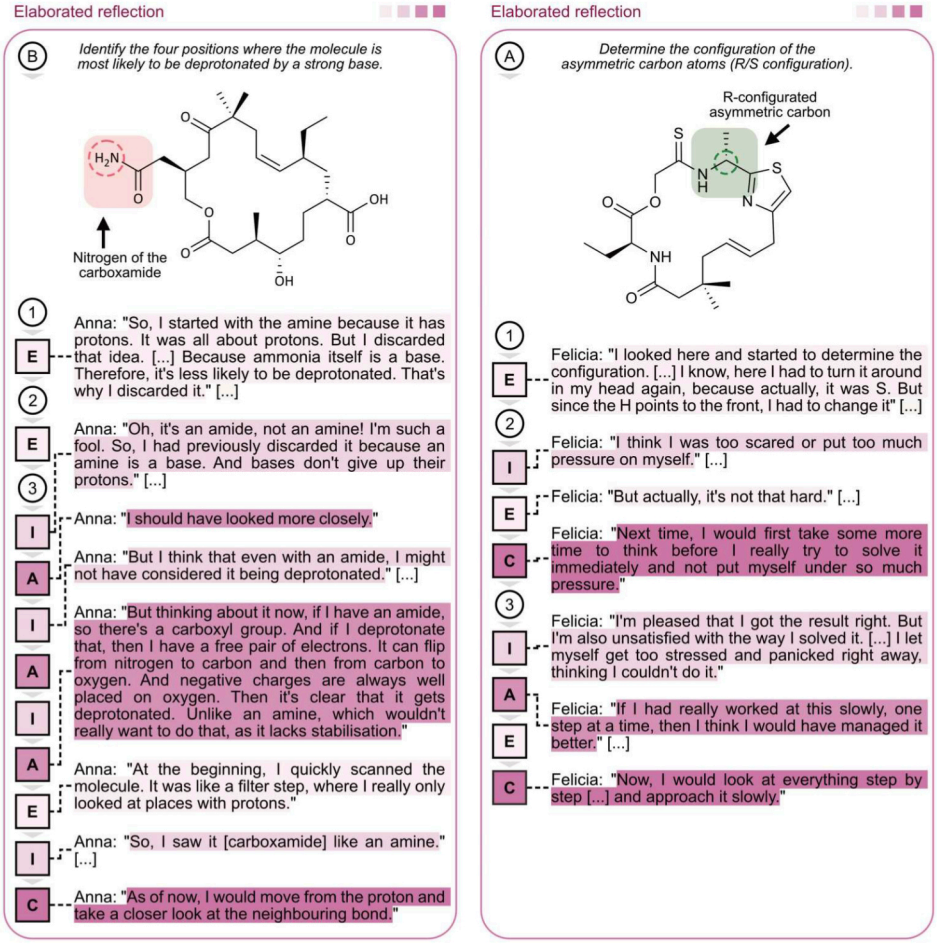


Figure 5. Excerpts of elaborated reflection patterns at the nitrogen of the carboxamide of the task about Brønsted acid-base concept (left, Anna's example) and at the R-configurated asymmetric carbon of the task about R/S configuration of asymmetric carbon atoms (right, Felicia's example).

centred reflection pattern, consequences were also not necessarily drawn subsequently after evaluating approaches. Thus, as the pattern emerged based on the reflection elements, it became apparent that evaluating approaches and drawing consequences are, thus, not necessarily connected. Additionally, by comparing the reflections of Peter, Susan, and Anna (Figures 4 and 5), it seems that in some cases, whether a purposeful consequence can be drawn relies on the students' ability to activate sufficient knowledge to make sense of the task solution. However, when consequences were drawn, the framing of the retrospective might have stimulated students to formulate more visual-related consequences.

In general, it appears that the accuracy of determining the structural site might influence the complexity of the reflection processes, which is further discussed in the next section.

3.2. To what extent are reflection patterns of higher complexity stimulated by certain phases along the eye-gaze-augmented retrospective and are dependent on student performance?

3.2.1. Change of the reflection patterns across the three phases of the eye-gaze-augmented retrospective

Changes in the reflection patterns across the three phases of the eye-gaze-augmented retrospective were quantified (Figure 6) to characterize how they stimulate reflection patterns of higher complexity.

In the first phase of the retrospective, reflection patterns were only stimulated by the eye-gaze replay and the prompt *What did you do and think?* Most reflection patterns in the first phase were reflection patterns of lower complexity (88%). Only a small proportion of reflection patterns showed a higher structural complexity (12%). In this phase, it was common for students to use their eye movements depicted in the eye-gaze replay to structure the description of their problem-solving strategy and redescribe their behavior. Students used the replay as a cue to follow events and recall their problem-solving. Students often started by describing their initial visual strategy (e.g. like in Felicia's reflection process in Figure 5) and, afterward, added their own (conceptual) thoughts and interpretations. This aligns with van Gog et al. (2005), who showed that dynamic eye-gaze replays can be used as memory cues for science problem-solving tasks. However, this phase seems to initiate a look back on action with limited incidences of reasoning through issues or consequences. Additionally, students often redescribed their visual problem-solving process, reiterating chronologically their problem-solving process. This calls for attention when using rather unguided eye-gaze replays in instruction, for instance, in future webcam-based eye-tracking scenarios.

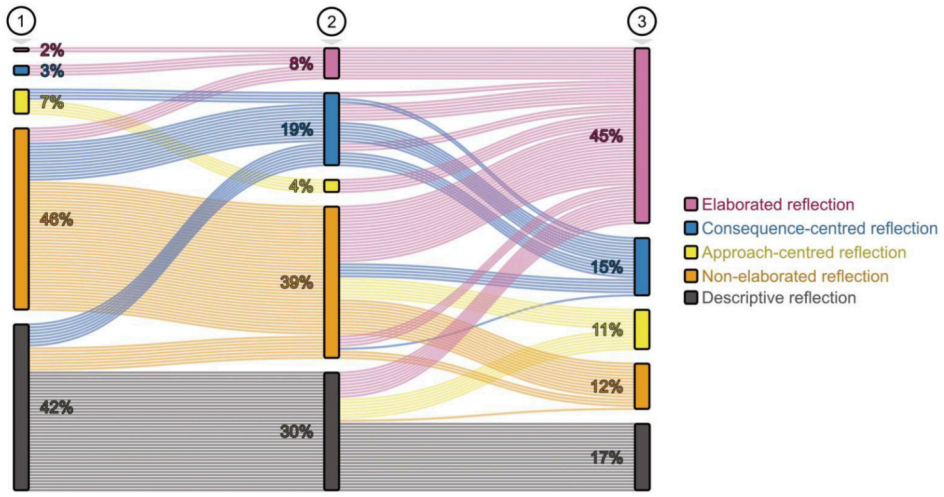


Figure 6. Occurrence of reflection patterns across the three phases of the eye-gaze-augmented retrospective. Each line depicts one reflection process throughout the retrospective at one structural site of the respective tasks. The numbers represent the percentage distribution of reflection patterns in the respective phase.

In the second phase of the retrospective, the complexity of patterns slightly increased, moving towards *consequence-centred reflections* (from 3% to 19%) and showing a decrease in reflection patterns with lower complexity (from 88% to 69%). Overall, there was only a slight shift in some reflection patterns during this phase. The additional prompts (see section 2.2.3.), at least the ones in the second phase of the retrospective, did not have a major impact compared to the think-aloud prompt and the eye-gaze replay in the first phase.

Only upon the introduction of the task solution with reflection prompts in the third phase of the retrospective did a major proportion of the reflection processes change. Some reflection processes changed from *descriptive reflection* to *elaborated reflection* (e.g. exemplified by Anna's reflection process in Figure 5). While *approach-centred reflection* and *consequence-centred reflection* constituted a minor portion of the reflection patterns throughout the eye-gaze-augmented retrospective (increasing slightly from 21% to 23%), *elaborated reflection* became predominant during the third phase (from 8% to 45%).

As the reflection patterns are based on the occurrence of reflection elements, it became evident that the first and third phases of the eye-gaze-augmented retrospective triggered different reflection elements. For instance, evaluating an approach was mostly observed in the third phase, as the given task solution confirmed the students' solutions or assumptions or revealed missed solutions. Nearly a third of the reflection patterns were of lower complexity at the end of the retrospective (29%).

Each of these three phases provided new information (e.g. task solution) or prompts that students could consider when reflecting on their problem-solving. However, that students were engaged in reflection processes of higher structural complexity is not an easy win, as reflection can depend on contextual factors (Boud et al., 1985), such as the accuracy of one's own solution.

3.2.2. Influence of students' accuracy

Figure 7 depicts the proportion of reflection processes based on the accuracy of determined structural sites.

Whether the structural site was determined accurately or inaccurately, the distribution of patterns of lower and higher complexity remained similar in the first two phases of the retrospective, with a predominant proportion of reflection patterns being of lower complexity. Only in the third phase of the retrospective did the distribution of patterns of lower and higher complexity differ remarkably.

In this phase, the dominant proportion of the reflection patterns on accurately determined structural sites was of lower complexity (63%). It seems that most of the students felt no need to reflect further, as the structural sites reflected on were already determined accurately. For *descriptive reflection*, this is plausible as students did not identify any issues (like Steven in his reflection in Figure 3). In contrast, in *non-elaborated reflection*, where students identified issues, they did not evaluate different approaches or draw consequences (like Gwen in her reflection in Figure 3). Most of the time, those students stopped their reflection as soon as they verified the accuracy of their response. Other students, nevertheless, tried to evaluate their approach and draw consequences, even though their correct response was already confirmed at the beginning of the third phase (as seen in Felicia's reflection in Figure 5).

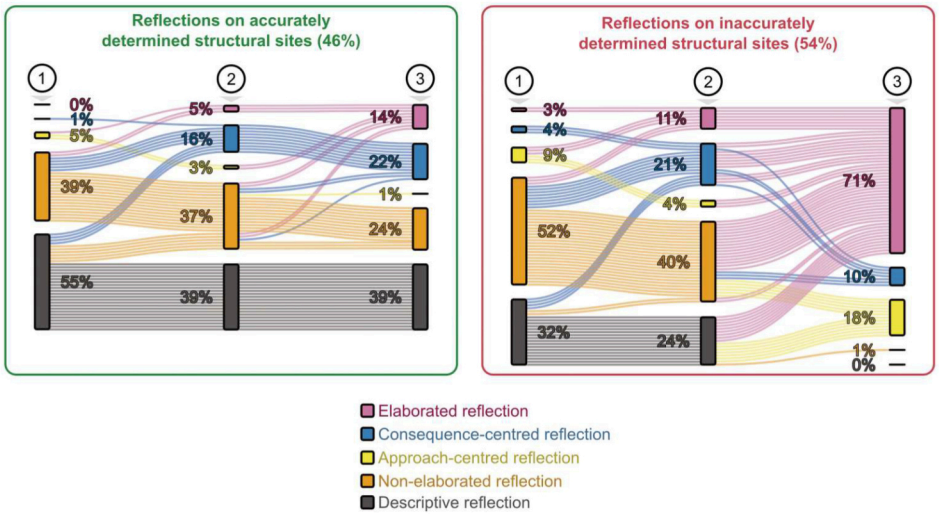


Figure 7. Occurrence of the reflection patterns across the three phases of the eye-gaze-augmented retrospective by accurately (left) or inaccurately (right) determined structural sites. 46% of the reflection processes occurred on accurately determined structural sites, while 54% reflection processes occurred on inaccurately determined structural sites. The numbers represent the percentage distribution of reflection patterns in the respective phase.

In comparison, almost every reflection process on inaccurately determined structural sites was of higher complexity in the third phase of the retrospective, with a predominant proportion of *elaborated reflection* (71% of 99%). In this phase, the students were made aware that they had inaccurately determined the structural site. It seems that realizing this when seeing the task solution prompted an enhanced reflection, leading students to evaluate different approaches, draw consequences, or do both.

In conclusion, students often used the presented task solution in the third phase of the retrospective to verify if the structural site was determined accurately or not. This is the reason why the structural complexity of most of the reflection patterns remained lower in the first and second phases of the retrospective and then either changed slightly if the structural site was determined accurately (Figure 7, left) or mostly changed to higher complexity if the structural site was determined inaccurately (Figure 7, right). Consequently, and not surprisingly, it seems that accurately determining the structural site did not foster more complex reflection patterns, as students often seemed pleased and saw no need to further reflect on a successful problem-solving process (e.g. Steven's and Gwen's example, Figure 3). Inaccurately determining a structural site may elicit feelings of inner discomfort, surprise, unwanted outcome, confusion, or difficulties in understanding, and provoke the intention to understand the content by themselves based on prior knowledge, which are all stimulators for reflection (Boyd & Fales, 1983; Moon, 2004; Schön, 1983) and may lead to reflection patterns of higher complexity.

3.3. Summary of results

To summarize, the eye-gaze replay and reflection prompts of the first two phases of the retrospective helped students structure their process of looking back on and being aware of essential aspects of their problem-solving process. The presented task solution, along with the accompanying reflection prompts, in the third phase of the retrospective had the greatest influence on the trajectory of the observed reflection patterns – especially when structural sites were incorrectly determined, reflection patterns of higher complexity were observable. It is, however, apparent that presenting the task solution in cases where structural sites were determined accurately only rarely supported the construction of reflection patterns of higher complexity with the current design of the retrospective. Even though the retrospective might be more beneficial for students who face challenges in solving problems, 71% of all analyzed reflection processes in this study were reflections of higher complexity, showing that the retrospective has the potential to stimulate the reflection processes of students with different needs and prior knowledge. The remaining 29% of reflection processes consisted almost exclusively of either *descriptive reflection* or *non-elaborated reflection* on accurately determined structural sites. Additional prompts may be necessary to enhance these reflection processes toward more complex patterns. For students who are confident and successful in their problem-solving approach (as illustrated by Steven's example in Figure 3), those prompts should reinforce productive behavior or student's metacognitive reflection. For students who face uncertainties or challenges yet achieve an accurate solution (as illustrated by Gwen's example in Figure 3), the prompts should encourage them to confront and resolve their uncertainties or issues. Purposeful prompting and guidance may be crucial, as students may not independently employ strategies to reflect on both the outcomes and the processes used in problem-solving for various reasons (e.g. feeling incapable of reflecting or considering it unnecessary) (Blackford et al., 2023).

3.4. Limitations

Some limitations must be considered in the presented study. Our analysis focused on single reflection processes on different structural sites in several tasks. While this approach provides insights into the range of reflection processes that occur in different task contexts, it limits the ability to compare and characterie individual students' approaches across all tasks.

Although the results indicate that an eye-gaze augmented retrospective is fruitful in engaging students in high-complexity reflection patterns, a comparison to a non-eye-gaze replay retrospective was not part of this study, leaving the specific benefits of eye-gaze replay unexplored. At the end of the data collection sessions, participants did report that the eye-gaze replay offered additional insights that might be inaccessible otherwise, indicating that cueing a retrospective with eye-gaze replays allows students to perceive their experience from another perspective.

Lastly, while our findings may imply that structurally more complex reflection patterns enhance learning outcomes (i.e. students' performance following the retrospective or enhanced productive visual behavior), learning outcomes were not a part of this analysis as we focused on reflection processes on structural sites and not on individual students. Hence, future investigations will include analysing eye-tracking data and students' performance.

4. Conclusion and implications for research and teaching

In this exploratory study, we investigated to what extent different phases of an eye-gaze-augmented retrospective can stimulate reflection processes on science problem-solving with representations by first characterizing the structural complexity of reflection processes on typical structural sites of organic-chemical representations in problem-solving tasks and second identifying the contribution of each phase of the eye-gaze-augmented retrospective to the reflection process while considering the accuracy of students' problem-solving.

Overall, the results suggest that an eye-gaze-augmented retrospective can be used to stimulate reflection processes on problem-solving with representations, but that specifically presenting the task solution, accompanied by reflection prompts, might increase reflection pattern complexity and is more suitable for students who struggle to solve the problem. In some cases, however, we observed that small-scale, conceptual support provided in a stepwise manner could have been helpful to support these students in formulating purposeful consequences when the content reflected on exceeds students' capabilities. This might be valuable when students cannot make sense of a task solution they have missed or have chosen incorrectly. Especially given that the task solution and the interview did not provide conceptual reasons.

Besides this, the results showed that reflecting in an eye-gaze-augmented retrospective on inaccurately solved problems might lead to productive learning situations. Inaccurately solved problems stem from various individual reasons, i.e. visual and conceptual causes, which become explicit through reflection in the retrospective. This allows students to identify and address the underlying reasons or challenges that led to the inaccuracies. Also, these insights can help instructors create adaptive support tailored to students' needs.

In addition, the designed eye-gaze-augmented retrospective mainly stimulates reflection processes of higher complexity on inaccurately solved problems – thus, the retrospective, especially for students with accurately determined solutions, has not yet tapped into the possible potential. If students solved problems accurately, that does not imply that these students did not benefit from reflection. Often, reflection is associated with learning from failures, but besides focusing on deficits and avoiding them, it may also be valuable for students to reinforce or refine productive behavior or deepen their understanding of the content. This type of reasoning, on the one hand, might not be captured by the typical categories used for reflection processes, which often focus on formulating alternatives or consequences. Future research should explore how retrospectives could involve a stronger focus on metacognition, e.g. reinforcing approaches in which students monitor or evaluate their problem-solving process, even when they are successful. On the other hand, one can assume that observing students ending the reflection process as soon as they have their accuracy confirmed might be related to typical epistemological messages on what knowledge products are valued (i.e. valuing authorised solutions) in the classroom, which might be more often conveyed in traditional courses (Russ, 2018).

Furthermore, the eye-gaze-augmented retrospective tended to stimulate students to focus more on their visual behavior and draw visual-related consequences, likely due to the visual framing as well as the visual stimuli of the retrospective. Hence, further

prompts may be necessary to stimulate students to focus additionally on the conceptual aspects linked to their visual behavior.

To conclude, even though this exploratory study may report first results, further research is necessary on how an eye-gaze-augmented retrospective should be designed (i.e. which additional prompts and scaffolds should be incorporated) and used (i.e. different problem-solving contexts and educational settings) to foster each student individually and more effectively in their problem-solving with representation. With these additional insights, such an eye-gaze-augmented retrospective could be integrated into a digital learning environment, providing timely feedback and scaffolding that stimulates students to reflect on their problem-solving process. Within such an online learning environment, differentiation could be conceivable by (1) offering tailored scaffolds based on students' accuracy, or (2) adapting prompts using machine learning to provide personalized feedback based on eye-tracking data, verbal data, or both.

Acknowledgements

We thank the German Research Foundation DFG for funding this research project. We also thank the students who voluntarily participated in the study, as well as Marietta Nied, Helena Rist, and Michael Schween for their support. This publication is part of the first author's doctoral (Dr rer. nat.) thesis in the Faculty of Biology and Chemistry at the Justus-Liebig-University Giessen, Germany.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the German Research Foundation DFG (Deutsche Forschungsgemeinschaft) (project number: 453926661).

Data availability statement

The participants of this study did not give written consent for their data to be shared. Therefore, supporting data is not available.

Ethics statement

All procedures performed in this study followed ethical standards (as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards applicable to educational research). While Institutional Review Board (IRB) or similar ethics approval is not mandated for chemistry education research at German universities, we adhered to a rigorous internal protocol to ensure ethical integrity. Participation in this study was entirely voluntary. Detailed informed written consent was obtained from all participants prior to data collection. This process involved providing each participant with comprehensive information regarding the study's objectives, participants' rights, the option to opt out at any time, and the data handling procedures. Consent forms

were signed and dated by each participant. No sufficiently identifiable information about participants is included in this manuscript.

ORCID

Axel Langner http://orcid.org/0000-0002-2274-8613
Nicole Graulich http://orcid.org/0000-0002-0444-8609

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, *48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712

Alpizar, D., Adesope, O. O., & Wong, R. M. (2020). A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development*, *68*(5), 2095–2119. https://doi.org/10.1007/s11423-020-09748-7

Baluyut, J. Y., & Holme, T. A. (2019). Eye tracking student strategies for solving stoichiometry problems involving particulate nature of matter diagrams. *Chemistry Teacher International*, *1*(1), 1–13. https://doi.org/10.1515/cti-2018-0003

Blackford, K. A., Greenbaum, J. C., Redkar, N. S., Gaillard, N. T., Helix, M. R., & Baranger, A. M. (2023). Metacognitive regulation in organic chemistry students: How and why students use metacognitive strategies when predicting reactivity. *Chemistry Education Research and Practice*, *24*(3), 828–851. https://doi.org/10.1039/D2RP00208F

Bollen, L., van Kampen, P., Baily, C., & De Cock, M. (2016). Qualitative investigation into students' use of divergence and curl in electromagnetism. *Physical Review Physics Education Research*, *12*(2), 020134. https://doi.org/10.1103/PhysRevPhysEducRes.12.020134

Bollen, L., van Kampen, P., Baily, C., Kelly, M., & De Cock, M. (2017). Student difficulties regarding symbolic and graphical representations of vector fields. *Physical Review Physics Education Research*, *13*(2), 020109. https://doi.org/10.1103/PhysRevPhysEducRes.13.020109

Bollen, L., van Kampen, P., & De Cock, M. (2015). Students' difficulties with vector calculus in electrodynamics. *Physical Review Special Topics - Physics Education Research*, *11*(2), 020129. https://doi.org/10.1103/PhysRevSTPER.11.020129

Boud, D., Keogh, R., & Walker, D. (1985). *Reflection: Turning experience into learning*. Routledge.

Boyd, E. M., & Fales, A. W. (1983). Reflective learning. *Journal of Humanistic Psychology*, *23*(2), 99–117. https://doi.org/10.1177/0022167883232011

Braun, I., Langner, A., & Graulich, N. (2022). Let's draw molecules: Students' sequential drawing processes of resonance structures in organic chemistry. *Frontiers in Education*, *7*, 1–24. https://doi.org/10.3389/feduc.2022.1055280

Brennan, R. L., & Prediger, D. J. (1981). Coefficient Kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, *41*(3), 687–699. https://doi.org/10.1177/001316448104100307

Cook, M., Wiebe, E. N., & Carter, G. (2008). The influence of prior knowledge on viewing and interpreting graphics with macroscopic and molecular representations. *Science Education*, *92*(5), 848–867. https://doi.org/10.1002/sce.20262

Cooper, M. M., Grove, N., Underwood, S. M., & Klymkowsky, M. W. (2010). Lost in Lewis structures: An investigation of student difficulties in developing representational competence. *Journal of Chemical Education*, *87*(8), 869–874. https://doi.org/10.1021/ed900004y

Corrales, K. A., & Erwin, T. C. (2020). Twitter and reflection: Tweeting towards deeper learning. *Reflective Practice*, *21*(4), 484–498. https://doi.org/10.1080/14623943.2020.1779048

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, *12*(1), 91–142. https://doi.org/10.1207/S15327809JLS1201_4

Dewey, J. (1933). *How we think*. Heath & Co Publishers.

Emhardt, S. N., Kok, E., van Gog, T., Brandt-Gruwel, S., van Marlen, T., & Jarodzka, H. (2023). Visualizing a task performer's gaze to foster observers' performance and learning—a systematic literature review on Eye movement modeling examples. *Educational Psychology Review*, *35*(1), 1–34. https://doi.org/10.1007/s10648-023-09731-7

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. The MIT Press.

Fleck, R., & Fitzpatrick, G. (2010). *Reflecting on reflection: Framing a design landscape*. OZCHI '10.

Frey, R. F., Brame, C. J., Fink, A., & Lemons, P. P. (2022). Teaching discipline-based problem solving. *CBE—Life Sciences Education*, *21*(2), 21:fe21. https://doi.org/10.1187/cbe.22-02-0030

Graulich, N., Hedtrich, S., & Harzenetter, R. (2019). Explicit *versus* implicit similarity – exploring relational conceptual understanding in organic chemistry. *Chemistry Education Research and Practice*, *20*(4), 924–936. https://doi.org/10.1039/C9RP00054B

Guo, L. (2022). How should reflection be supported in higher education? — A meta-analysis of reflection interventions. *Reflective Practice*, *23*(1), 118–146. https://doi.org/10.1080/14623943.2021.1995856

Halverson, K. L., Pires, C. J., & Abell, S. K. (2011). Exploring the complexity of tree thinking expertise in an undergraduate systematics course. *Science Education*, *95*(5), 794–823. https://doi.org/10.1002/sce.20436

Hansen, S. J. R., Hu, B., Riedlova, D., Kelly, R. M., Akaygun, S., & Villalta-Cerdas, A. (2019). Critical consumption of chemistry visuals: Eye tracking structured variation and visual feedback of redox and precipitation reactions. *Chemistry Education Research and Practice*, *20*(4), 837–850. https://doi.org/10.1039/C9RP00015A

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. https://doi.org/10.1016/0742-051X(94)00012-U

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K.-J., & Lehtinen, M. (2008). Gaze path stimulation in retrospective think-aloud. *Journal of Eye Movement Research*, *2*(4), 1–18. https://doi.org/10.16910/jemr.2.4.5

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., & Gerjets, P. (2013). Learning to see: Guiding students' attention via a model's eye movements fosters learning. *Learning and Instruction*, *25*, 62–70. https://doi.org/10.1016/j.learninstruc.2012.11.004

Killion, J. P., & Todnem, G. R. (1991). A process for personal theory building. *Educational Leadership*, *48*(6), 14–16.

King, A. (1994). Guiding knowledge construction in the classroom: Effects of teaching children how to question and how to explain. *American Educational Research Journal*, *31*(2), 338–368. https://doi.org/10.3102/00028312031002338

Klein, P., Viiri, J., Mozaffari, S., Dengel, A., & Kuhn, J. (2018). Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots? *Physical Review Physics Education Research*, *14*(1), 010116. https://doi.org/10.1103/PhysRevPhysEducRes.14.010116

Kolb, D. A., & Fry, R. E. (1975). Toward an applied theory of experiential learning. In C. Cooper (Ed.), *Theories of group process* (pp. 33–57). John Wiley & Sons.

Korthagen, F. A. (1985). Reflective teaching and preservice teacher education in The Netherlands. *Journal of Teacher Education*, *36*(5), 11–15. https://doi.org/10.1177/002248718503600502

Korthagen, F. A., & Kessels, J. P. (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational Researcher*, *28*(4), 4–17. https://doi.org/10.3102/0013189X028004004

Kozma, R. B., & Russell, J. (2005). Students becoming chemists: Developing representationl competence. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 121-145). Springer Netherlands. https://doi.org/10.1007/1-4020-3613-2_8

Kranz, D., Schween, M., & Graulich, N. (2023). Patterns of reasoning – exploring the interplay of students' work with a scaffold and their conceptual knowledge in organic chemistry. *Chemistry Education Research and Practice*, *24*(2), 453–477. https://doi.org/10.1039/D2RP00132B

Kuckartz, U., & Rädiker, S. (2019). *Analyzing qualitative data with MAXQDA*. Springer.

Marshall, T., Keville, S., Cain, A., & Adler, J. R. (2021). On being open-minded, wholehearted, and responsible: A review and synthesis exploring factors enabling practitioner development in reflective practice. *Reflective Practice*, *22*(6), 860–876. https://doi.org/10.1080/14623943.2021.1976131

McClary, L., & Talanquer, V. (2011). Heuristic reasoning in chemistry: Making decisions about acid strength. *International Journal of Science Education*, *33*(10), 1433–1454. https://doi.org/10.1080/09500693.2010.528463

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice*. Psychology Press.

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176–1189. https://doi.org/10.1111/medu.12583

Prain, V., & Tytler, R. (2012). Learning through constructing representations in science: A framework of representational construction affordances. *International Journal of Science Education*, *34*(17), 2751–2773. https://doi.org/10.1080/09500693.2011.626462

Rodemer, M., Eckhard, J., Graulich, N., & Bernholt, S. (2020). Decoding case comparisons in organic chemistry: Eye-tracking students' visual behavior. *Journal of Chemical Education*, *97*(10), 3530–3539. https://doi.org/10.1021/acs.jchemed.0c00418

Rodemer, M., Eckhard, J., Graulich, N., & Bernholt, S. (2021). Connecting explanations to representations: Benefits of highlighting techniques in tutorial videos on students' learning in organic chemistry. *International Journal of Science Education*, *43*(17), 2707–2728. https://doi.org/10.1080/09500693.2021.1985743

Rodemer, M., Lindner, M. A., Eckhard, J., Graulich, N., & Bernholt, S. (2022). Dynamic signals in instructional videos support students to navigate through complex representations: An eye-tracking study. *Applied Cognitive Psychology*, *36*(4), 852–863. https://doi.org/10.1002/acp.3973

Russ, R. S. (2018). Characterizing teacher attention to student thinking: A role for epistemological messages. *Journal of Research in Science Teaching*, *55*(1), 94–120. https://doi.org/10.1002/tea.21414

Sandars, J. (2009). The use of reflection in medical education: AMEE guide No. 44. *Medical Teacher*, *31*(8), 685-695. https://doi.org/10.1080/01421590903050374

Schön, D. A. (1983). *The reflective practitioner*. Temple Smith.

Schönborn, K. J., & Anderson, T. R. (2009). A model of factors determining students' ability to interpret external representations in biochemistry. *International Journal of Science Education*, *31*(2), 193–232. https://doi.org/10.1080/09500690701670535

Shemshack, A., Kinshuk, & Spector, J. M. (2021). A comprehensive analysis of personalized learning components. *Journal of Computers in Education*, *8*(4), 485-503. https://doi.org/10.1007/s40692-021-00188-7

Susac, A., Planinic, M., Bubic, A., Jelicic, K., & Palmovic, M. (2023). Linking information from multiple representations: An eye-tracking study. *Frontiers in Education*, *8*, 1–10. https://doi.org/10.3389/feduc.2023.1141896

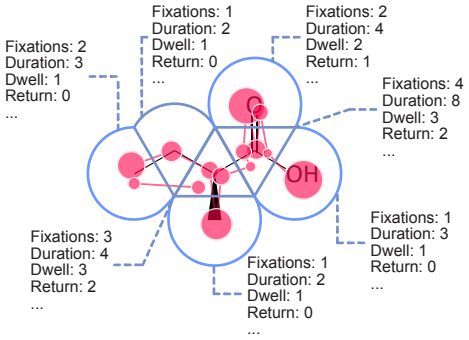
Tóthová, M., & Rusek, M. (2022). "Do you just have to know that?" novice and experts' procedure when solving science problem tasks. *Frontiers in Education*, *7*, 1–14. https://doi.org/10.3389/feduc.2022.1051098

Tsai, P.-Y., Yang, T.-T., She, H.-C., & Chen, S.-C. (2019). Leveraging college students' scientific evidence-based reasoning performance with Eye-tracking-supported metacognition. *Journal of Science Education and Technology*, *28*(6), 613–627. https://doi.org/10.1007/s10956-019-09791-x

van Gog, T., Paas, F., Van Merriënboer, J. J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237. https://doi.org/10.1037/1076-898X.11.4.237

van Gog, T., & Rummel, N. (2010). Example-Based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, *22*(2), 155–174. https://doi.org/10.1007/s10648-010-9134-7

DEFINING AREAS OF INTEREST



Langner, A., Hain, L., & Graulich, N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, 102(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830

JOURNAL OF CHEMICAL EDUCATION

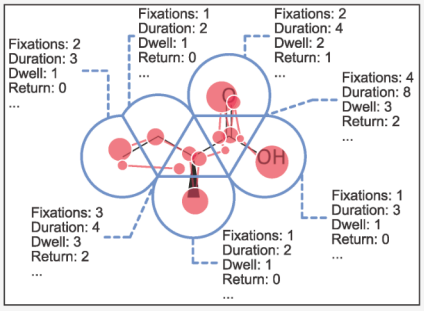
pubs.acs.org/jchemeduc
Communication

Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research

Axel Langner, Lea Sophie Hain, and Nicole Graulich*

Cite This: J. Chem. Educ. 2025, 102, 1285−1297
Read Online

ABSTRACT: Often, eye-tracking researchers define areas of interest (AOIs) to analyze eye-tracking data. Although AOIs can be defined with systematic methods, researchers in organic chemistry education eye-tracking research often define them manually, as the semantic composition of the stimulus must be considered. Still, defining appropriate AOIs during data preparation is crucial, as it can impact AOI-based metrics—even to the extent of their significance. However, defining AOIs can be challenging when dealing with complex and dense stimuli, such as organic representations. To address this challenge, we introduce an already established systematic method, the limited-radius Voronoi tessellation (LRVT) method, which has not yet been used in chemistry education research. While being algorithmic, this method still considers the semantic composition of the stimulus and could offer an alternative to manually defining AOIs in organic representations. By applying the LRVT method, the position of atoms can be used to construct AOIs that represent the closest area to these atoms within a defined radius. In this communication, we illustrate the how-to of constructing AOIs with the LRVT method during data preparation and demonstrate how these AOIs can be used to analyze and compare visual problem-solving behavior in different grain sizes through exemplary analyses. Finally, we discuss the advantages and limitations of this alternative method.



KEYWORDS: *Graduate Research, Organic Chemistry, Problem Solving, Assessment, Chirality*

INTRODUCTION

Eye tracking is a valuable tool for investigating visual attention distribution and how a learner processes representations. This makes it particularly useful in chemistry education research, where analyzing learners' use of representations, formulas, models, and diagrams is important. Consequently, in the last decades, several eye-tracking chemistry education research studies have focused on areas such as problem-solving strategies,[1−4] expert-novice comparison,[5−8] students' use of representations,[9−12] evaluation of learning material,[13−19] cognitive demand,[20−23] predictability of eye movements,[24,25] or feedback.[26,27]

While analyzing eye-tracking data, the researchers define areas of interest (AOIs) – definite regions in the presented stimulus in which they aim to analyze eye-gaze data.[28] Thus, the setup of AOIs depends on the researcher's goal. For example, Braun et al.[4] investigated students' drawing processes of organic chemistry resonance structures. The researchers defined square-shaped AOIs at different levels of granularity (i.e., grain sizes). Fine-grain AOIs covered detailed features of single molecules (e.g., charges, structural sites) to investigate what structural features students attend to while drawing resonance structures. Coarse-grain AOIs covered broad elements, like textboxes, reactions, and entire drawings, to explore what elements students connect while constructing resonance structures. The shape, size, and position of AOIs are crucial, as AOI-based metrics like fixation duration, AOI hits, or transitions are used to draw conclusions about participants' viewing processes.[28,29] Thus, how AOIs should be set raises many questions: What grain size should be selected? What elements of the stimulus should be covered by AOIs? How close should AOIs be located to each other? How big should the margins of the AOIs be? How should the AOIs be shaped? How is AOI setup replicability ensured? Consequently, defining AOIs can be particularly challenging, especially in complex stimuli with multiple closely arranged elements, such as in organic representations—including Lewis structures, Haworth projections, chair conformations, wedge-and-dash projections, or entire reaction mechanisms. AOIs can be defined with several approaches: 1.) Researcher-defined AOIs are manually set by the researchers themselves. 2.) Stimulus-defined AOIs are created based on specific stimulus features detected by computational methods like edge detection or color segmenta-

Received: July 5, 2024
Revised: January 21, 2025
Accepted: January 24, 2025
Published: February 3, 2025

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

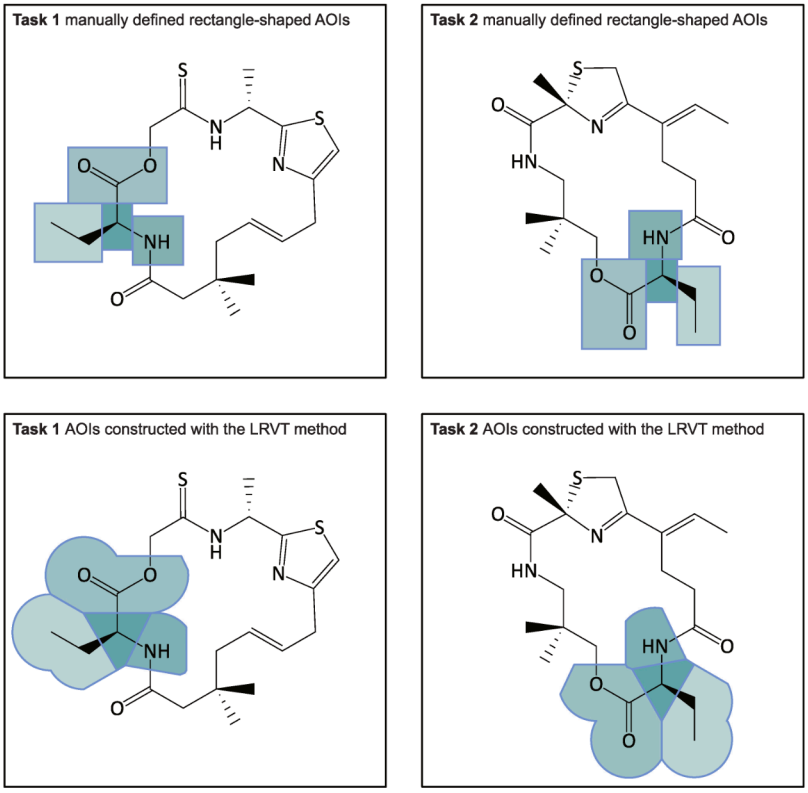


Figure 1. Comparison of researcher-defined AOIs (top) and LRVT-defined AOIs (bottom) across two different structures (left vs right).

tion. 3.) Algorithm-defined AOIs are calculated post hoc solely based on the collected gaze data using specific algorithmic methods such as clustering. 4.) Grid-based AOIs are constructed by placing a grid content-independently across the stimulus leading to each grid cell being an individual AOI.[28,30] However, it is often crucial to consider the semantic composition of the stimulus, which inevitably leads to researcher-defined AOIs. Hence, it is not surprising that organic chemistry eye-tracking researchers often define AOIs manually.[4−6,11,13,31] Although researchers define AOIs with their best intent and expertise, defining AOIs manually may introduce subjectivity and risk compromising comparability and replicability, especially when visual behavior is compared across multiple structures in a study.

In light of this, the limited-radius Voronoi tessellation (LRVT) method—an established systematic method for defining AOIs based on the spatial arrangement of elements in the stimulus—may assist organic chemistry eye-tracking researchers during data preparation in defining AOIs while considering the semantic composition of the representation. Hence, LRVT-defined AOIs could be particularly useful in instances when comparing eye movements across multiple structures or contexts. For example, researchers may consider comparing the eye movements of participants as they determine the configuration of an asymmetric carbon atom in two different structures. In this case, fine-grain AOIs could be constructed where each AOI encloses only one of the relevant features (i.e., the carbon center and atoms of the substituents). Rectangle-shaped AOIs could be defined manually (see Figure 1, top), but it may be challenging to define objective and resembling AOIs across stimuli that cover a sufficient area, including the white space between close elements, while also ensuring their size is appropriately defined and proportionally distributed to correspond to the spatial arrangement and number of those elements, in order to register relevant fixations and correctly assign them to the corresponding features in both structures—especially in dense stimuli. In such a case, the LRVT method offers the possibility to systematically define similar AOIs across both structures (see Figure 1, bottom). This ensures that any changes in eye movements are not artificially introduced by the set of the AOI itself.

To illustrate this method in this communication, we first describe how to apply the LRVT method to organic stimuli. Second, we illustrate its application with an example in which participants solved tasks involving organic representations. Finally, we discuss the method's advantages and limitations.

METHOD: THE LRVT METHOD FOR ORGANIC STIMULI

The LRVT method is based on the tessellation principle introduced by George Voronoi.[32] It divides an area into smaller areas, known as Voronoi cells, across multiple points called cell centers and limits the size of these cells by a set radius. Each cell represents the closest area to its cell center within a defined radius.

For organic stimuli, AOIs can be defined with the LRVT method by using atoms as cell centers. The method to construct

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

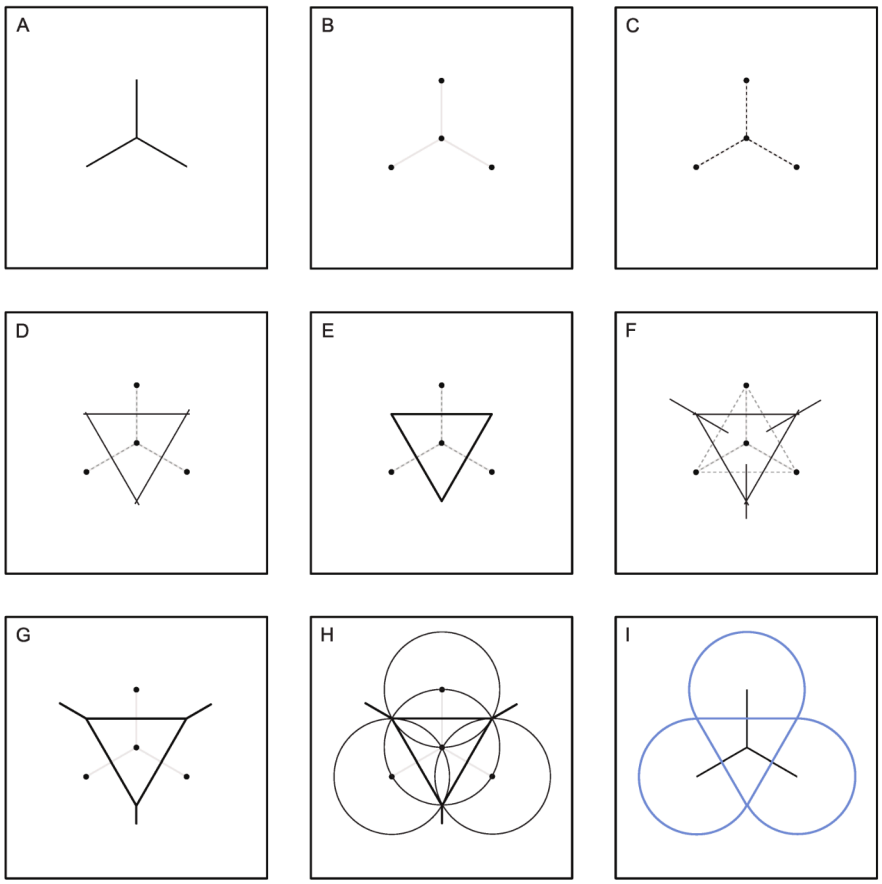


Figure 2. (A) Defining AOIs for an isobutane molecule representation following the LRVT method. (B) Each atom represents a cell center. (C) First, to determine the AOI of the tertiary carbon, draw connecting lines to atoms nearby. (D) Next, draw perpendicular bisecting lines for each connecting line. (E) The area enclosed by the perpendicular lines represents a Voronoi cell. (F) Repeat the steps for the other atoms. (G) The results are the corresponding Voronoi cells. (H) Limit each cell by a set radius. (I) Finally, the intersectional areas construct the LRVT-defined AOIs.

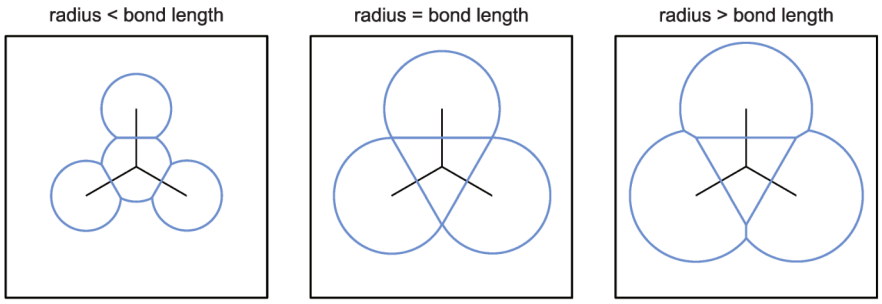


Figure 3. LRVT-defined AOIs with different radii.

LRVT-defined AOIs is demonstrated with an isobutane molecule representation in Figure 2. For a practical guide on constructing LRVT-defined AOIs for an organic representation, please refer to the Supporting Information.

With this method, fixations located in an LRVT-defined AOI can be assigned to the corresponding atom since these fixations fall in the area closest to this atom within a specific radius. The example in Figure 2I shows LRVT-defined AOIs with a radius equal to the bond length of the molecule. However, other radii can also be used (see Figure 3).

As shown in Figure 3, the radius defines the size and shape of the AOIs and is, therefore, an important variable. Since there are no standardized guidelines for setting AOI sizes, some eye-tracking researchers suggest considering several factors: the

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

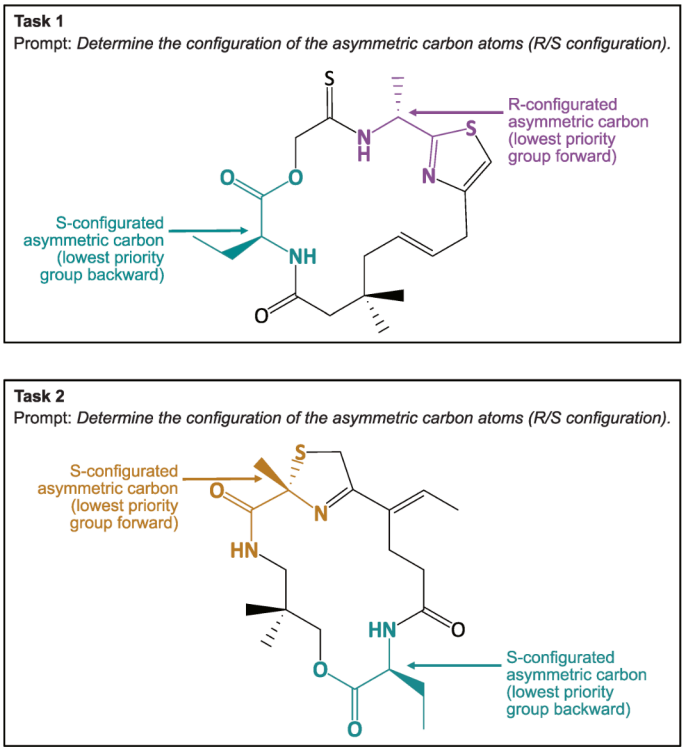


Figure 4. Tasks with task prompt and a large cyclic complex molecule. The two asymmetric carbon atoms in each task are highlighted. The asymmetric carbons highlighted in teal are identical. Solutions were hidden from the students.

accuracy and precision of the eye-tracking data, the importance of registering all the fixations, the risk of inflating the false negative or false positive rate, theoretical expectations about the distribution of fixations, the fixation variance across participants, the amount of white space, the element density of the stimulus, and the conspicuity of elements.[28,29,33,34] As a result, we recommend that AOIs for organic stimuli in chemistry education research, in general, should be defined larger in size for the following reasons: 1.) a significant amount of white space with areas of high element density is a typical characteristic of stimuli in organic chemistry, 2.) organic representations often include conspicuous elements (e.g., protruding peripheral atoms or functional groups) which can be visible at larger eccentricities,[34] 3.) it may be important to register every fixation, and 4.) the accuracy of the eye-tracking data is often not sufficient enough leading to spread fixations. To define adequately large AOIs, the radius in LRVT-defined AOIs should be set to the size of the bond length or larger.

A standardized radius is preferable for defining AOIs in a comparable manner. One fixed, researcher-independent, and easy-to-determine length in every organic stimulus is the bond length within a molecule. Therefore, setting the radius to the bond length of the molecules in the presented stimulus could guide the definition of AOIs.

Setting the radius to the bond length of the molecule provides an additional advantage. To register every relevant fixation, large AOI margins may be used.[28] Yet, using large margins may increase the false positive rate or result in overlapping AOIs in dense stimuli.[29] However, the LRVT method with a set radius to the bond length effectively addresses these issues. First, overlapping AOIs are avoided, as the AOI borders are always the perpendicular bisecting lines of connection between atoms (see Figure 2G). Second, as the radius of the LRVT-defined AOIs is set equal to the bond length of the depicted molecule the AOI margins—and consequently the AOI size—will dynamically adapt with the molecule's presentation size. To outline, if the molecule is presented at a large size no fixation distribution overlap can be expected due to a larger distance between elements. In this case, large AOI margins are recommended to include all nearby fixations, reducing the false negative risk.[29] Conversely, if the molecule is presented at a small size, fixation distribution overlap can be expected due to a reduced distance between elements. In this case, small AOI margins are recommended to balance the false negative and positive risk.[29] By dynamically adjusting the AOI size, the LRVT method with a radius set to the bond length ensures that both scenarios are effectively addressed.

In the next section, we apply the LRVT method with a defined radius equal to the bond length in a stimulus depicting a large complex molecule and illustrate how the method can be used to characterize a student's visual problem-solving behavior.

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

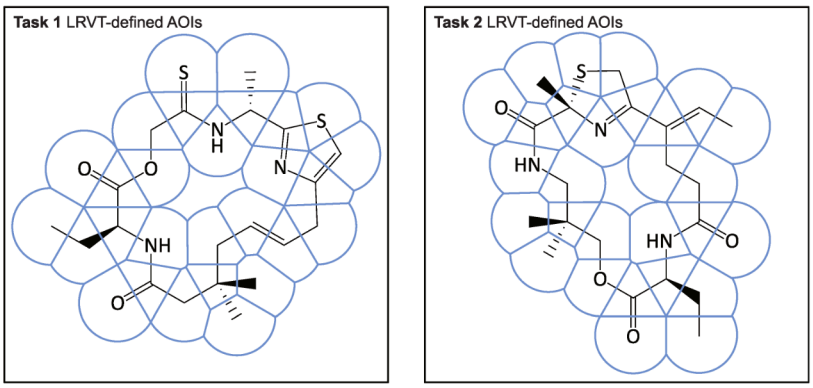


Figure 5. LRVT-defined AOIs.

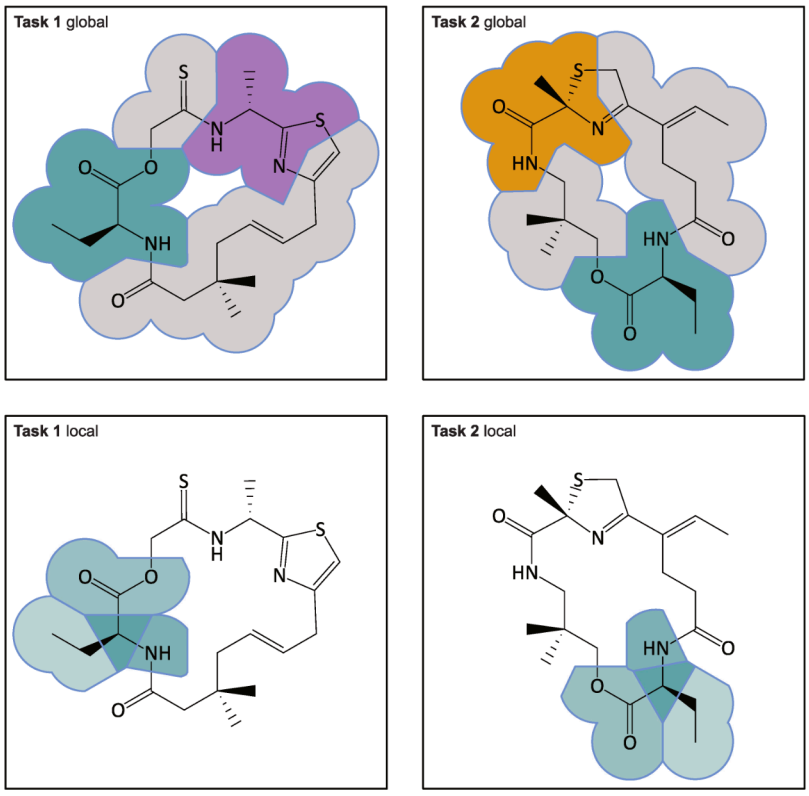


Figure 6. LRVT-defined AOIs in different grain sizes (global and local) for each task. Coarse-grain size, defining AOIs for broad structural features (i.e., global) (top). Fine-grain size, defining fine-grain AOIs for the asymmetric carbon centers and its adjacent atoms with the lowest priority group backward (i.e., local) (bottom). Color scheme: teal AOIs = asymmetric carbon atoms with the lowest priority group backward in task 1 and task 2, purple AOI = asymmetric carbon atom with the lowest priority group forward in task 1, ocher AOI = asymmetric carbon atom with the lowest priority group forward in task 2, grey AOIs = molecule chain.

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

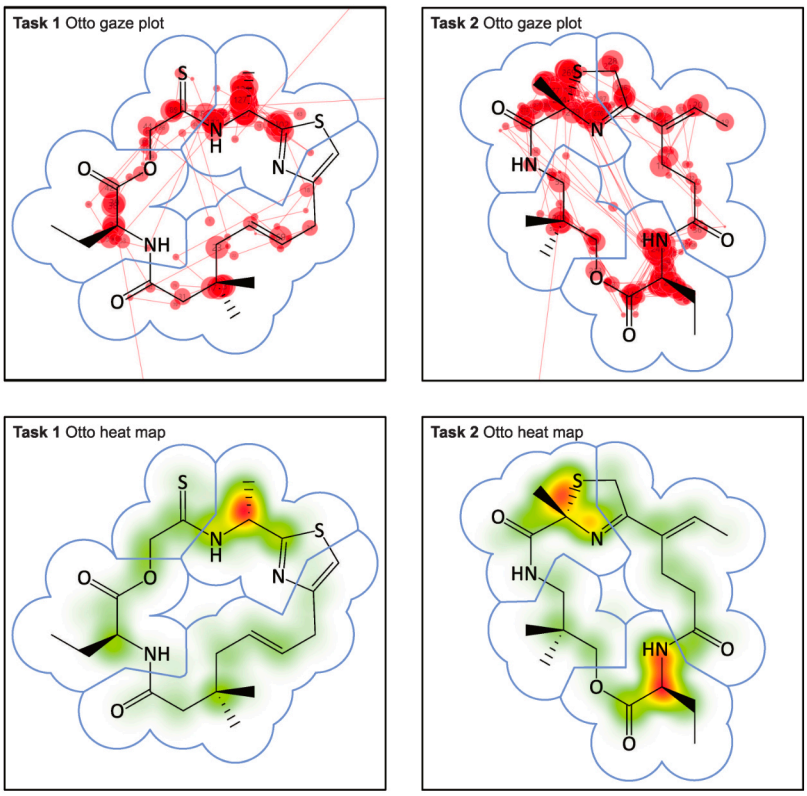


Figure 7. Gaze plots (top) and heat maps (bottom) of Otto's eye gaze while determining the configuration of asymmetric carbon atoms in the molecules of task 1 and task 2. The blue lines indicate the coarse-grain AOIs depicted in Figure 6, top. Note: The color intensity of the heatmaps reflects the relative fixation duration within each task and is not directly comparable between the two tasks, as the color scales are specific for each problem-solving process.

■ APPLICATION: DEMONSTRATION OF THE LRVT METHOD

Setting

In the following, eye-tracking recordings of two students solving an organic chemistry problem will be used to demonstrate the application of the LRVT method for data preparation.

During data collection with the eye tracker, for each task, participants were shown a single complex organic molecule and prompted to determine the configuration of asymmetric carbon atoms. Those large cyclic organic molecules are depicted in the wedge-and-dash projection and contain several functional groups and two asymmetric carbon atoms (see Figure 4). Originally, those tasks were used in a study about an eye-gaze-augmented retrospective.[35]

The eye-tracking data were collected with a sampling rate of 120 Hz with a Tobii Pro X3-120 screen-based eye-tracking system, a Tobii Pro External Processing Unit, and the associated software Tobii Pro Lab[36] using a live viewer setup with a 24-in. presentation monitor with a 1920 × 1080 resolution. The participants were instructed to keep a distance of 60 to 65 cm from the presentation monitor and to minimize head and body movements during the task to reduce data loss. Before starting each task, the system was calibrated using a nine-point calibration and a four-point validation. Immediately after the calibration, the participants were presented with the task prompt *Determine the configuration of the asymmetric carbon atoms (R/S configuration)*. They then continued by pressing a key, and a fixation cross was shown to the participants for two seconds before one of the molecules was automatically presented. The molecules were scaled so that the distance between atoms was 2.2 cm, which corresponds to an eye movement of approximately 2°. Therefore, this exceeds the accuracy of the eye-tracking system of 0.7° and the suggested 1−1.5° by Holmqvist et al.[28] After completing the task, the participants pressed a key to stop the eye-tracking data collection. Immediately upon completing each task, participants recorded their answers on a sheet of paper.

Defining AOIs in Complex Molecules

To determine the Voronoi cell centers for the molecules, the x and y coordinates of the atoms were extracted from the ChemDraw file of the molecules. The LRVT-defined AOIs were computed using the software R[37] with the R-packages deldir,[38] ggforce[39] and ggplot2.[40] The radius of the AOIs was set to the size of the bond length. The resulting R-plots were then imported into the eye-tracking software, and AOIs were drawn according to the calculated limited-radius Voronoi cells (see Figure 5). A detailed guide for this procedure is provided in the Supporting Information.

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

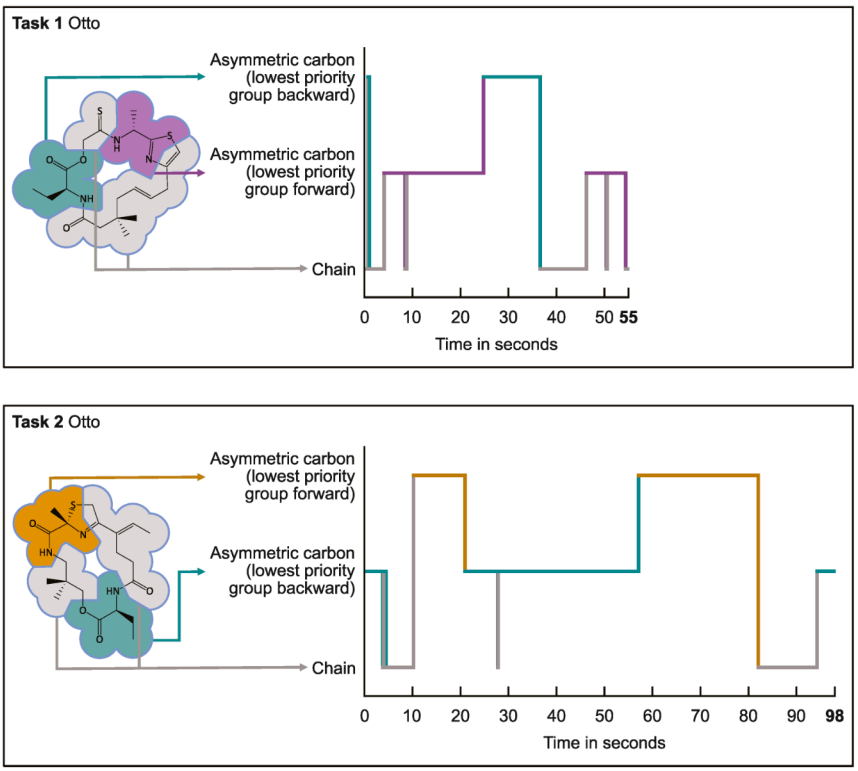


Figure 8. Time-sequence plots of Otto's eye gaze for both tasks. The colors correspond to the colors of the AOIs depicted in Figure 6, top. The color of each vertical line indicates the AOI from which the transition to the next AOI originates.

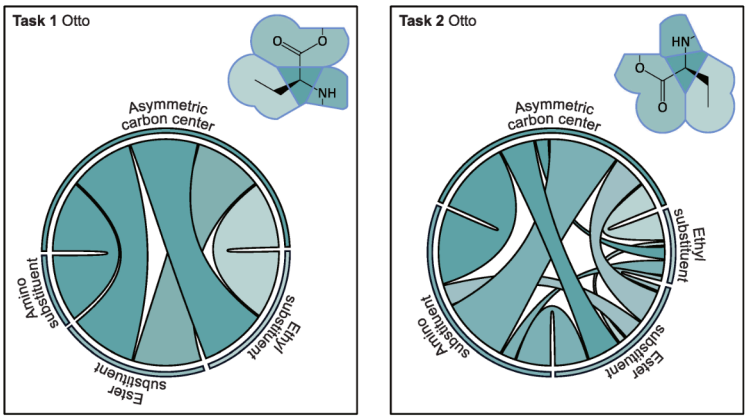


Figure 9. Relative number of transitions of Otto's eye gaze in both tasks between the fine-grain AOIs. The colors correspond to the colors of the AOIs depicted in Figure 6, bottom. The color of each chord indicates the AOI from which the transition originates.

With this, AOIs can be merged to construct AOIs of broader grain sizes, namely coarse-grain AOIs (see Figure 6, top) that group multiple elements together, as well as fine-grain AOIs (see Figure 6, bottom), where each AOI reflects distinct individual elements. This flexibility allows researchers to analyze specific AOIs either individually (e.g., the total fixation duration on the asymmetric carbon and its substituents as a whole group; see Figure 6, top) or as groups (e.g., the distribution of the fixation duration across the asymmetric carbon and its substituents; see Figure 6, bottom).

Using LRVT-Defined AOIs and Different Grain Sizes to Compare a Student's Visual Problem-Solving Behavior Across Representations

The eye-gaze data from an exemplary student's problem-solving processes is used to illustrate how the LRVT-defined AOIs (see

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

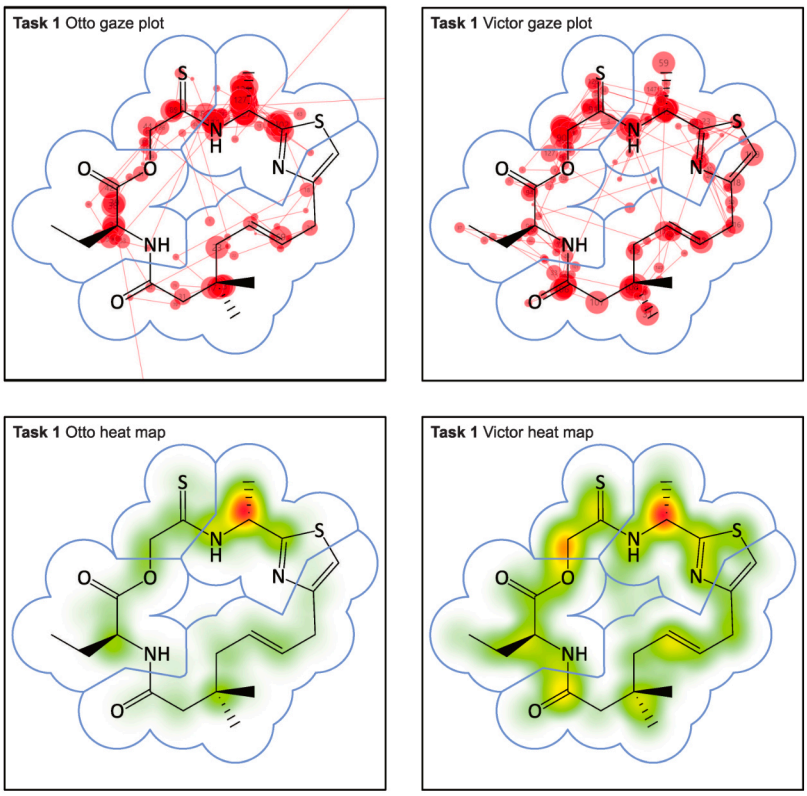


Figure 10. Gaze plots (top) and heat maps (bottom) of Otto's and Victor's eye gaze while determining the configuration of asymmetric carbon atoms in the molecule of task 1. The blue lines indicate the coarse-grain AOIs depicted in Figure 6, top. Note: The color intensity of the heatmaps reflects the relative fixation duration within each problem-solving process and is not directly comparable between the two students, as the color scales are specific for each problem-solving process.

Figure 6) can help to characterize and compare these processes, especially across tasks that have comparable stimuli. Otto, an organic chemistry beginner, identified and determined both asymmetric carbons accurately in both tasks (see Figure 4). Otto's eye gaze along the presented molecules is depicted in Figure 7.

As illustrated in Figure 7, the LRVT-defined AOIs encompass a sufficient area of the stimuli, ensuring that all of Otto's fixations on the molecules are registered. Additionally, the LRVT-defined coarse-grain AOIs (see Figure 6, top) appropriately partition the stimuli, assigning Otto's fixations to the corresponding features. This allows us to globally analyze his consecutive eye movements and the absolute time he spent on the coarse-grain AOIs, as visualized in the time-sequence plots (Figure 8)

When comparing the absolute time Otto spent on the respective coarse-grain AOIs in Figure 8, it becomes evident that Otto spent most time looking at the asymmetric carbons in both tasks. Contrastingly, Otto spent less time looking at the chain and more time on the asymmetric carbon atom with the lowest priority group backward in task 2 than in task 1, even though they are structurally identical.

The fine-grain AOIs (see Figure 6, bottom) enable a more detailed analysis of Otto's gaze on these asymmetric carbon atoms locally to analyze his process of determining their configuration and identifying differences in his gaze. As these AOIs are set similarly and cover comparable areas across both asymmetric carbon atoms, this preserves consistency and prevents the introduction of artificial variations in the eye movements by the AOIs themselves. Taking the relative number of transitions between the fine-grain AOIs of each asymmetric carbon into account, Otto performed more transitions within the asymmetric carbon atom in task 2 (see Figure 9, right) than in task 1 (see Figure 9, left).

The systematic definition of AOIs using the LRVT method within similarly designed tasks allowed us to characterize Otto's visual behavior and to identify differences across both tasks. We can infer that, globally, Otto spent most of his time determining the configuration of the two asymmetric carbon atoms in each task (see Figures 7 and 8). Locally on the asymmetric carbons with the lowest priority group backward, his eye movements often shifted between the carbon centers and adjacent groups. It could be that the change in the representation of the identical asymmetric carbon atom with the lowest priority group backward resulted in more transitions between the AOIs (see Figure 9). Thus, by using the LRVT-defined AOIs at different grain sizes, subtle differences in a student's visual problem-solving behavior across representations with dense features, like asymmetric carbon atoms, become comparable and visible, as they ensure that all fixations are registered and assigned to the

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

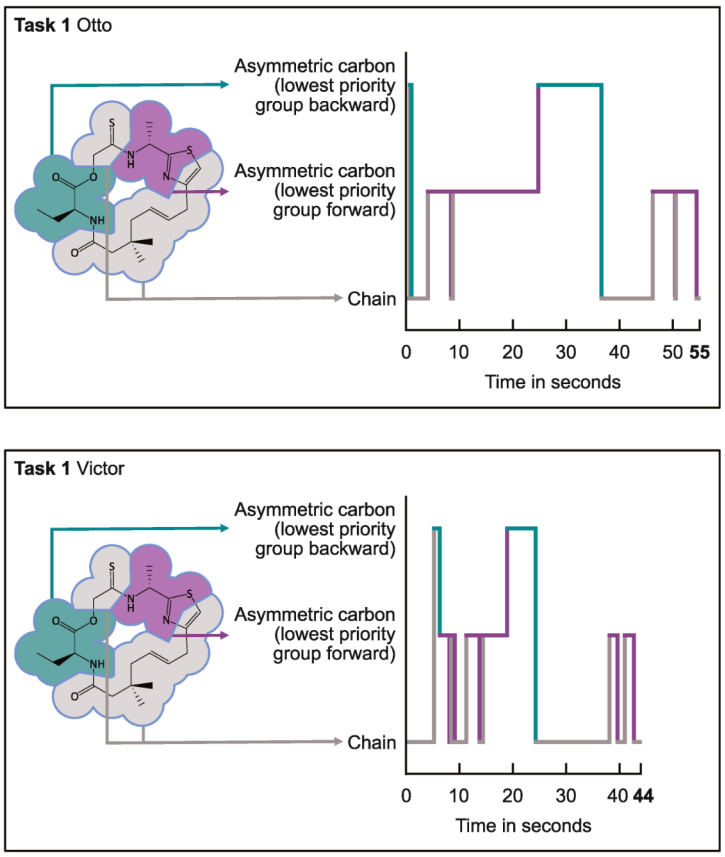


Figure 11. Time-sequence plots of Otto's and Victor's eye gaze for tasks 1. The colors correspond to the colors of the AOIs depicted in Figure 6, top. The color of each vertical line indicates the AOI from which the transition to the next AOI originates.

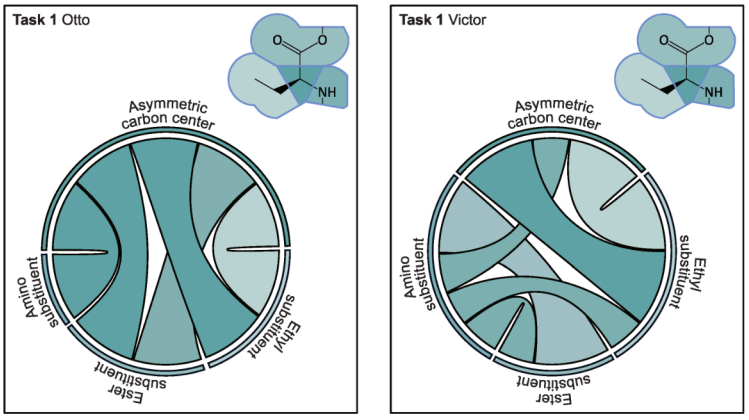


Figure 12. Relative number of transitions of Otto's and Victor's eye gaze in task 1 between the fine-grain AOIs. The colors correspond to the colors of the AOIs depicted in Figure 6, bottom. The color of each chord indicates the AOI from which the transition originates.

corresponding feature, maintain consistency in size and shape across representations.

Using LRVT-Defined AOIs and Different Grain Sizes to Compare Visual Problem-Solving Behavior Across Students

In addition to using the LRVT method to prepare eye-tracking data to enhance the comparability of students' visual problem-

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

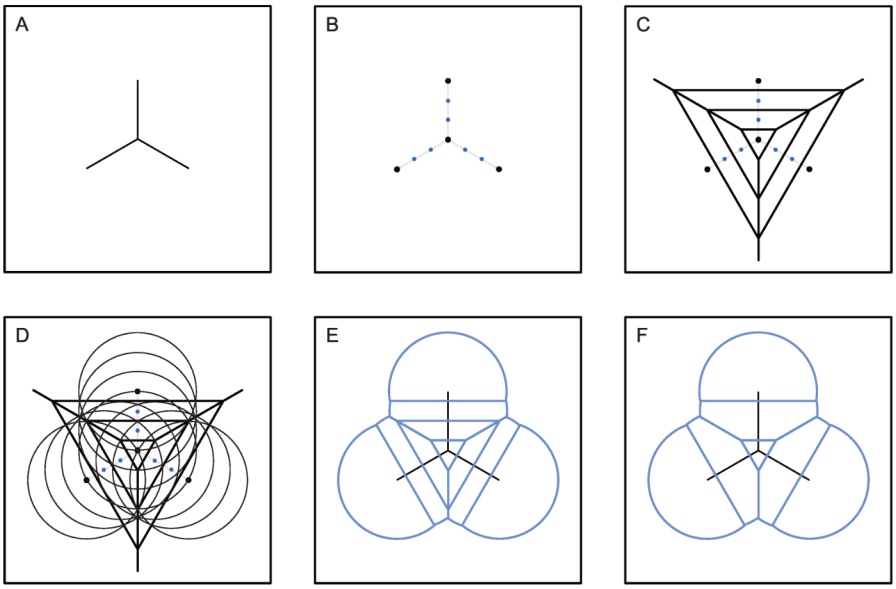


Figure 13. (A) Defining AOIs for each atom and bond of an isobutane molecule following the LRVT method. (B) Each atom and two points of each bond represent a cell center. (C) Voronoi cells are created for each cell center. (D) Limit each cell by a radius equaling the bond length. (E) The intersectional areas represent the LRVT-defined AOIs. (F) Constructing the final AOIs for each bond by merging the involved LRVT-defined AOIs.

solving across representations, the LRVT-defined AOIs may also be used to compare visual problem-solving behavior across students. This is because the LRVT method covers the entire area of structural features, eliminating gaps between AOIs and ensuring that every fixation within these areas is registered to reveal subtle differences in eye movements of different students.

To exemplify this comparison across students, Otto's and Victor's visual problem-solving behavior of task 1 can be compared. Victor, another organic chemistry beginner, only determined the configuration of the asymmetric carbon with the lowest priority group forward, unlike Otto, who determined the configuration of both asymmetric carbons. Both Otto's and Victor's eye gaze of task 1 along the presented molecules are depicted in Figure 10.

As mentioned before, the LRVT-defined AOIs encompass a sufficient area of the stimuli. While comparing the gaze of different students, this is particularly important given the spread of fixations caused by fixation variance across participants. This becomes apparent by comparing the gaze plots of Otto's and Victor's gaze behavior (see Figure 10, top). Even though Otto's and Victor's fixations are distributed differently (i.e., Otto's fixations are closer to the representation while Victor's fixations are slightly more spread out), the LRVT-defined AOIs ensure that every fixation is registered. We are, thus, able to globally analyze their consecutive eye movements and the absolute time they spent on the coarse-grain AOIs, as visualized in the time-sequence plots (Figure 11).

Interestingly, even though Otto and Victor differed in their performance (Otto accurately determined the configuration of the asymmetric carbon with the lowest priority group backward, while Victor did not), they processed the molecule in a similar pattern on a global level (see Figure 11). The fine-grain AOIs can be used to analyze their local visual problem-solving behavior on the asymmetric carbon atom in relation to their performance. Considering the relative number of transitions between the fine-grain AOIs, Victor's gaze behavior is characterized by a higher variety of different types of transitions (Figure 12, right) compared to Otto's gaze behavior (Figure 12, left).

Using the LRVT-defined AOI-based analysis, we can infer that Victor demonstrated a comparable pattern to Otto globally (see Figure 11), even though Victor did not determine the configuration of one of the asymmetric carbon atoms. Locally, Victor showed differences in the variety of transitions on this asymmetric carbon atom (see Figure 12).

With this example, we demonstrate that LRVT-defined AOIs can be used to identify differences in visual problem-solving behavior across students. All fixations of the participants, despite fixation variance, are registered within those dense areas to reveal subtle differences in eye movements.

■ DISCUSSION: APPLICABILITY OF THE LRVT METHOD

With Otto's and Victor's gaze behavior in task 1 and task 2, we presented two example scenarios of using the LRVT method for eye-tracking data preparation. Nonetheless, despite its demonstrated advantages, there are a few limitations to consider while using LRVT-defined AOIs. As the construction of the Voronoi cells is based on zero-dimensional elements, i.e., points, the method is not directly applicable to one-dimensional elements (e.g., lines, arrows) or two-dimensional elements (e.g., rectangles, circles). A workaround might be to translate the multidimensional elements of the stimulus into individual points by converting the lines into points with an appropriate distance and then constructing one merged AOI based on the calculated Voronoi cells. With this approach, it would be possible to calculate AOIs for the lines that depict bonds in a molecule (see

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

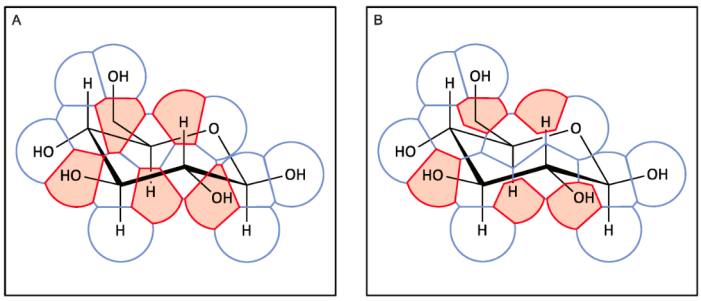


Figure 14. (A) LRVT-defined AOIs for D-glucose in its β-pyranose form in chair conformation. Red AOIs overlap with bonds between other atoms. (B) Manual adjustment of the red AOIs to avoid overlapping with bonds between other atoms.

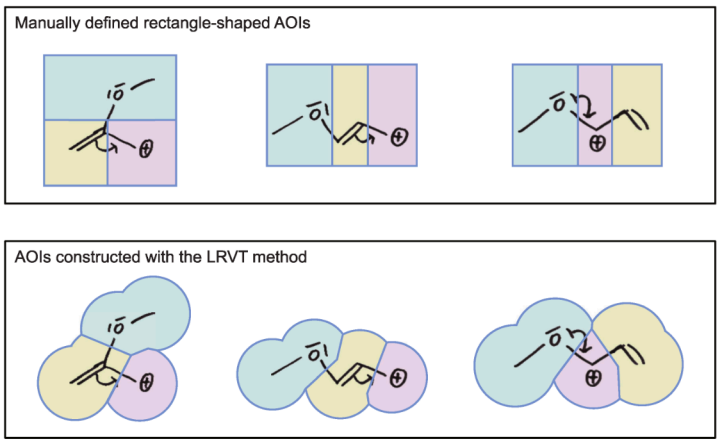


Figure 15. Comparison of researcher-defined AOIs for the main features of the drawn structures (methoxy group, double bend, and positive charge) used in the study by Braun et al.[4] (top) and the LRVT-defined AOIs for these respective main features (bottom). Since bond lengths can vary within a drawing, the mean bond length of each drawing was used as the radius for the LRVT method.

Figure 13). This could be crucial in contexts in which the type of bond might be important (e.g., resonance).

Additionally, with this method, it is possible to construct AOIs at a very fine-grain size, leading to an extensive number of smaller AOIs. This would lead to a lower total fixation count and duration per AOI and more entries and transitions, making the data challenging to interpret or potentially not useful for deriving teaching implications. Hence, we recommend that eye-tracking researchers select purposeful AOI-grain sizes depending on the research question by constructing LRVT-defined AOIs based on the atoms and merging the very fine-grain AOIs into AOIs with broader grain sizes for functional groups or whole molecules (exemplified in Figure 6).

Furthermore, we propose setting the radius in the LRVT method equal to the bond size of the molecule. Ideally, the accuracy of the eye tracker is considered during the experiment's design. However, we acknowledge that with current technology, it is sometimes impossible to depict a whole reaction mechanism, arrange the molecules at a proper distance, and scale the molecules so that the bond length exceeds the accuracy of the eye tracker. If the radius falls below the accuracy of the eye tracker, two options should be considered. First, one solution could be to set the radius in the LRVT method to at least match the accuracy of the eye tracker, ensuring that scattered fixations—due to eye-tracker inaccuracy—are still assigned to an AOI. This works well for cases where no other AOIs are closely positioned, e.g., peripheral AOIs (see Figure 3). Second, in more complex cases with closely positioned AOIs and fixation distribution overlap, it may be necessary to merge AOIs to a broader grain size to prevent the misattribution of fixations.

Finally, using the LRVT method in three-dimensional organic representations, such as Haworth projections, chair conformations, or Newman projections, may lead to improper AOIs. In these cases, it may be necessary to evaluate the LRVT-defined AOIs and to adjust them manually. For example, when constructing AOIs using the LRVT method for D-glucose in its β-pyranose form in chair conformation, some AOIs of atoms overlap with the bonds between other atoms (see Figure 14A). However, if a fixation falls on a bond, it should be assigned to at least one of the atoms involved. Therefore, these LRVT-defined AOIs may require manual adjustment, making them more similar to researcher-defined AOIs (see Figure 14B).

■ CONCLUSION

In this communication, we aimed to inform the research community about a potentially practical method for organic chemistry eye-tracking research, which has been used in other areas of eye-tracking research outside science education.[41,42] We introduced the LRVT method for defining AOIs in organic stimuli as an alternative to other AOI-defining approaches,

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

especially researcher-defined AOIs, during data preparation, as it also considers the semantic composition. Furthermore, we demonstrated how to apply this method in eye-tracking data preparation and illustrated through examples how these AOIs in two different grain sizes can be used to characterize visual problem-solving behavior across structural representations and students systematically.

Despite its limitations, this method has several advantages. As a systematic method, it solves many challenges researchers regularly face when defining AOIs (e.g., choosing the size, shape, and location or balancing the false positive and false negative rate).[28,29,41] Additionally, it is based on mathematical operations, and therefore, the AOIs can be calculated quickly with software like R or similar tools (an R script to calculate the LRVT cells is provided in the Supporting Information). Moreover, since the AOIs are defined around the position of the atoms, this method considers the semantic composition of the molecule, even though the LRVT method calculates AOIs algorithmically. Finally, by using a fixed radius equaling the bond length, the LRVT method provides a more objective way to define AOIs and avoids inflating the false negative or false positive rate, as the size of the AOI is dynamically adjusted to the stimulus properties.

With this, the method may improve the replicability of eye-tracking analysis and the comparability of results across stimuli and contexts. Nevertheless, we want to emphasize that dependent on the stimuli and how manually defined AOIs are set, using the LRVT method in data preparation may not per se lead to differences in eye-tracking data processing (i.e., resulting metrics). Defining AOIs manually is still a purposeful method. For instance, in one of our eye-tracking studies,[4] we manually defined fine-grain AOIs on given structures and students' drawings to investigate students' sequential drawing processes of resonance structures (see Figure 15, top). This allowed us to analyze students' fixation duration on the features of these structures to draw conclusions about how students integrate information during their drawing process. However, since manually defining AOIs carries the risk of subjectivity—particularly when comparing eye movements across multiple structures or contexts—the LRVT method can offer valuable support in maintaining consistency in AOI definitions—even in students' drawings (see Figure 15, bottom). While using the researcher-defined AOIs in our study does not undermine the findings and conclusions, defining AOIs with the LRVT method could have assisted us in retrospect during the data preparation process and may have enhanced the replicability and comparability.

Broadly speaking, the challenges the organic chemistry education eye-tracking research community faces are similar to those other behavioral eye-tracking research communities encounter (e.g., defining AOIs, software limitations, insufficient data quality, conducting naturalistic real-world studies).[41,43−46] To address both these discipline-specific and broader challenges and to advance eye tracking as a research method, the community should explore and evaluate more robust and systematic approaches, such as the LRVT method, to ensure reliable and meaningful insights.

For now, within the context of organic chemistry education eye-tracking research, the LRVT method, despite its limitations, offers a practical approach for supporting researchers in eye-tracking data preparation. It allows for the definition of AOIs at atom-level granularity in dense stimuli, offering flexibility in selecting the grain size by merging fine-grain AOIs. Due to its systematic nature, it enables comparisons of eye movements across different structures or contexts.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.4c00830.

Step-by-step instructions on how to construct LRVT-defined AOIs from a molecule's ChemDraw file using R (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

Authors

Axel Langner − *Institute of Chemistry Education, Justus-Liebig-University, 35392 Giessen, Germany;* orcid.org/0000-0002-2274-8613

Lea Sophie Hain − *Institute of Chemistry Education, Justus-Liebig-University, 35392 Giessen, Germany;* orcid.org/0009-0005-5200-0832

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.4c00830

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the German Research Foundation (DFG) for funding this research project (project number: 453926661), Timo Dahms for his support in data preparation, and Zoie Bunch for her continuous support.

■ REFERENCES

(1) Baluyut, J. Y.; Holme, T. A. Eye tracking student strategies for solving stoichiometry problems involving particulate nature of matter diagrams. *Chem. Teach. Int.* **2019**, *1* (1), 1−13.

(2) Cullipher, S.; Sevian, H. Atoms versus bonds: how students look at spectra. *J. Chem. Educ.* **2015**, *92* (12), 1996−2005.

(3) Tang, H.; Kirk, J.; Pienta, N. J. Investigating the effect of complexity factors in stoichiometry problems using logistic regression and eye tracking. *J. Chem. Educ.* **2014**, *91* (7), 969−975.

(4) Braun, I.; Langner, A.; Graulich, N. Let's draw molecules: Students' sequential drawing processes of resonance structures in organic chemistry. *Front. Educ.* **2022**, *7*, 1−24.

(5) Connor, M. C.; Glass, B. H.; Finkenstaedt-Quinn, S. A.; Shultz, G. V. Developing Expertise in ^{1}H NMR Spectral Interpretation. *J. Org. Chem.* **2021**, *86* (2), 1385−1395.

(6) Rodemer, M.; Eckhard, J.; Graulich, N.; Bernholt, S. Decoding Case Comparisons in Organic Chemistry: Eye-Tracking Students' Visual Behavior. *J. Chem. Educ.* **2020**, *97* (10), 3530−3539.

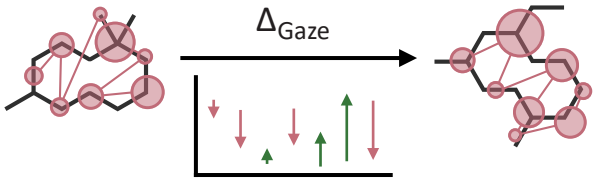
(7) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR spectra through the eyes of a student: eye tracking applied to NMR items. *J. Chem. Educ.* **2017**, *94* (1), 29−37.

(8) Tóthová, M.; Rusek, M. Do you just have to know that?" Novice and experts' procedure when solving science problem tasks. *Front. Educ.* **2022**, *7*, 1051098.

(9) Hinze, S. R.; Rapp, D. N.; Williamson, V. M.; Shultz, M. J.; Deslongchamps, G.; Williamson, K. C. Beyond ball-and-stick: Students' processing of novel STEM visualizations. *Learn. Instr.* **2013**, *26*, 12−21.
(10) Tóthová, M.; Rusek, M.; Chytrý, V. Students' procedure when solving problem tasks based on the periodic table: An eye-tracking study. *J. Chem. Educ.* **2021**, *98* (6), 1831−1840.
(11) Williamson, V. M.; Hegarty, M.; Deslongchamps, G.; Williamson, K. C., III; Shultz, M. J. Identifying student use of ball-and-stick images versus electrostatic potential map images via eye tracking. *J. Chem. Educ.* **2013**, *90* (2), 159−164.
(12) Locatelli, S. W.; de Oliveira, L. S.; da Silva S, R., Jr.; Sato, J. R. Perspectives of the Eye-Tracking Application to Understand Access to Representational Levels in Chemistry. *J. Sci. Educ.Technol.* **2024**, *33*, 823.
(13) Rodemer, M.; Lindner, M. A.; Eckhard, J.; Graulich, N.; Bernholt, S. Dynamic signals in instructional videos support students to navigate through complex representations: An eye-tracking study. *Appl. Cognit. Psychol.* **2022**, *36* (4), 852−863.
(14) VandenPlas, J. R.; Herrington, D. G.; Shrode, A. D.; Sweeder, R. D. Use of simulations and screencasts to increase student understanding of energy concepts in bonding. *J. Chem. Educ.* **2021**, *98* (3), 730−744.
(15) Nehring, A.; Busch, S. Setting Up Chemistry Demonstrations According to the Left-to-Right Principle: An Eye-Movement-Pattern-Based Analysis. *J. Chem. Educ.* **2020**, *97* (1), 275−282.
(16) Fleischer, T.; Deibl, I.; Moser, S.; Strahl, A.; Maier, S.; Zumbach, J. Mobile Eye Tracking during Experimenting with Digital Scaffolding—Gaze Shifts between Augmented Reality and Experiment during Zinc Iodide Electrolysis Set-Up. *In Educ. Sci.* **2023**, *13*, 170.
(17) Huangfu, Q.; Li, H.; Tang, S.; Wang, J.; Liu, Q.; Chen, G. How teacher enthusiasm affects students' learning of chemistry declarative knowledge in video lectures. *Chem. Educ. Res. Pract.* **2022**, *23* (4), 898−912.
(18) Syskowski, S.; Huwer, J. A Combination of Real-World Experiments and Augmented Reality When Learning about the States of Wax. An Eye-Tracking Study. *Educ. Sci.* **2023**, *13*, 177.
(19) Yang, F.-Y.; Wang, H.-Y. Tracking visual attention during learning of complex science concepts with augmented 3D visualizations. *Comput. Educ.* **2023**, *193*, No. 104659.
(20) Rodemer, M.; Karch, J.; Bernholt, S. Pupil dilation as cognitive load measure in instructional videos on complex chemical representations. *Front. Educ.* **2023**, *8*, 1062053.
(21) Karch, J. M.; García Valles, J. C.; Sevian, H. Looking into the black box: Using gaze and pupillometric data to probe how cognitive load changes with mental tasks. *J. Chem. Educ.* **2019**, *96* (5), 830−840.
(22) Tang, H.; Pienta, N. Eye-tracking study of complexity in gas law problems. *J. Chem. Educ.* **2012**, *89* (8), 988−994.
(23) Ye, J.; Zheng, Y.; Zhan, M.; Zhou, Y.; Li, L.; Chen, D. Characteristics of Pre-Service Chemistry Teachers' Mechanistic Reasoning In Organic Chemistry Tasks: An Eye-Tracking Study. *Res. Sci. Educ.* **2024**.
(24) Chen, S.-C.; Hsiao, M.-S.; She, H.-C. The effects of static versus dynamic 3D representations on 10th grade students' atomic orbital mental model construction: Evidence from eye movement behaviors. *Comput. Hum. Behav.* **2015**, *53*, 169−180.
(25) Ghali, R.; Ouellet, S.; Frasson, C. LewiSpace: An Exploratory Study with a Machine Learning Model in an Educational Game. *J. Educ. Train. Stud.* **2015**, *4* (1), 192−201.
(26) Langner, A.; Graulich, N.; Nied, M. Eye-Tracking as a Promising Tool in Pre-Service Teacher Education—A New Approach to Promote Skills for Digital Multimedia Design. *J. Chem. Educ.* **2022**, *99* (4), 1651−1659.
(27) Hansen, S. J. R.; Hu, B.; Riedlova, D.; Kelly, R. M.; Akaygun, S.; Villalta-Cerdas, A. Critical consumption of chemistry visuals: eye tracking structured variation and visual feedback of redox and precipitation reactions. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 837−850.
(28) Holmqvist, K.; Nyström, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. *Eye Tracking: A Comprehensive Guide to Methods and Measures*; Oxford University Press, 2011.
(29) Orquin, J. L.; Ashby, N. J. S.; Clarke, A. D. F. Areas of Interest as a Signal Detection Problem in Behavioral Eye-Tracking Research. *Journal of Behavioral Decision Making* **2016**, *29* (2−3), 103−115.
(30) Fuhl, W.; Kuebler, T.; Santini, T.; Kasneci, E. Automatic generation of saliency-based areas of interest for the visualization and analysis of eye-tracking data. *Proceedings of the Conference on Vision, Modeling, and Visualization* **2018**.
(31) Weinrich, M.; Britt, R. Students' Attention on Curved Arrows While Evaluating the Plausibility of an Organic Mechanistic Step. In *Student Reasoning in Organic Chemistry*, Graulich, N., Shultz, G., Eds.; The Royal Society of Chemistry, 2022; pp 3−18.
(32) Voronoi, G. Nouvelles applications des paramètres continus à théorie des formes quadratiques. Deuxième Mémoire. Recherches sur les parallélòèdres primitifs. *J. Reine Angew. Math.* **1909**, *134*, 198−287.
(33) Goldberg, J. H.; Helfman, J. I. Comparing information graphics: a critical look at eye tracking. *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization* **2010**, 71−78.
(34) Hooge, I. T.; Camps, G. Scan path entropy and arrow plots: capturing scanning behavior of multiple observers. *Front. Psychol.* **2013**, *4*, 00996.
(35) Langner, A.; Graulich, N. From Sight to Insight − Reflection Processes in an Eye-Gaze-Augmented Retrospective. *Int. J. Sci. Educ.* **2024**, 2430804.
(36) *Tobii Pro Lab*; Tobii AB: Stockholm, 2024. https://connect.tobii.com/s/lab-downloads?language=en_US (accessed Jul 2024).
(37) *R: A language and environment for statistical computing*; *R Foundation for Statistical Computing*: Vienna, Austria, 2021. https://www.R-project.org/ (accessed May 2022).
(38) deldir: Delaunay Triangulation and Dirichlet (Voronoi) Tessellation. *Comprehensive R Archive Network (CRAN)*, 2021. https://cran.r-project.org/package=deldir (accessed May 2022).
(39) *ggforce: Accelerating 'ggplot2'*, 2024. https://ggforce.data-imaginist.com (accessed May 2022).
(40) *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphic*; Springer-Verlag: New York, 2016. https://ggplot2.tidyverse.org (accessed May 2022).
(41) Hessels, R. S.; Kemner, C.; van den Boomen, C.; Hooge, I. T. C. The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods* **2016**, *48* (4), 1694−1712.
(42) de Winter, J. C. F.; Dodou, D.; Tabone, W. How do people distribute their attention while observing The Night Watch? *Perception* **2022**, *51* (11), 763−788.
(43) Vehlen, A.; Standard, W.; Domes, G. How to choose the size of facial areas of interest in interactive eye tracking. *PLoS One* **2022**, *17* (2), No. e0263594.
(44) Orquin, J. L.; Holmqvist, K. Threats to the validity of eye-movement research in psychology. *Behavior Research Methods* **2018**, *50* (4), 1645−1656.
(45) Jarodzka, H.; Holmqvist, K.; Gruber, H. Eye tracking in Educational Science: Theoretical frameworks and research agendas. *J. Eye Mov Res.* **2017**.
(46) Lappi, O. Eye Tracking in the Wild: the Good, the Bad and the Ugly. *Journal of Eye Movement Research* **2015**.

https://doi.org/10.1021/acs.jchemed.4c00830
J. Chem. Educ. 2025, 102, 1285−1297

CHANGES IN EYE MOVEMENTS



Langner, A., Sahba, M., Popova, M., & Graulich, N. (2025). An Integrated Approach to Characterizing Changes in Organic Chemistry Students' Eye Movements. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

JOURNAL OF CHEMICAL EDUCATION

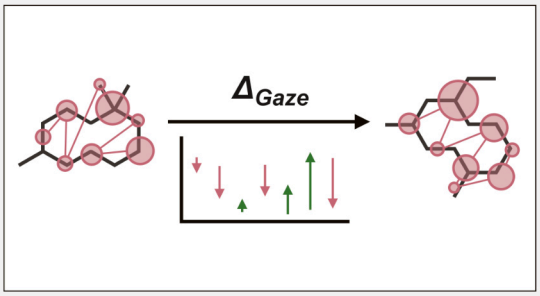
pubs.acs.org/jchemeduc
Chemical Education Research

An Integrated Approach to Characterizing Changes in Organic Chemistry Students' Eye Movements

Axel Langner, Marie Sahba, Maia Popova, and Nicole Graulich*

Cite This: https://doi.org/10.1021/acs.jchemed.5c00849
Read Online

ABSTRACT: Although the use of representations is crucial for problem-solving in chemistry, students often encounter challenges when using them. Hence, various interventions have been designed to support students' use of chemistry representations, which can motivate students to modify their visual behavior. However, to fully understand how these interventions affect students, it is crucial not only to understand how they are using representations but also to understand how and to what extent their visual behavior changes in response to these interventions. Since learning may induce complex changes in eye movement across spatial, temporal, and psychophysiological dimensions, a key methodological challenge is how to comprehensively characterize these changes in instructional settings. However, relying on single metrics offers only fragmented insights, capturing isolated aspects of students' information processing, whereas simultaneous interpretation of multiple eye-tracking metrics presents considerable challenges. In light of this, we conducted two studies using stereochemistry tasks to explore how changes in eye movements can be comprehensively characterized in students who participated in an intervention and those who did not. By calculating the average absolute bounded normalized change across multiple eye-tracking metrics, as Δ_{Gaze}, and additionally examining the change of individual eye-tracking metrics across tasks, we were able to characterize both overall and detailed changes in students' eye movements while considering the multifaceted nature of eye-tracking data. In this article, we illustrate the potential and limitations of this methodological approach.



KEYWORDS: *Graduate Research, Chemical Education Research, Organic Chemistry, Problem Solving, Assessment, Chirality*

INTRODUCTION

In recent years, eye-tracking technology has proven to offer unique insights into students information processing in education research.[1−3] Especially, in chemistry, a domain characterized by complex visual representations such as structures, formulas, and diagrams, eye-tracking provides access to students' problem-solving processes by recording spatial, temporal, and psychophysiological information on their visual behavior. Prior research has shown that different levels of expertise substantially guide the use of representations in problem-solving. With increasing expertise, less time is taken, relevant information is focused more selectively, and a greater variety of information is integrated.[4−10] For example, Connor et al.[5] compared the visual behavior of undergraduate and doctoral students while interpreting ^{1}H NMR spectra. While doctoral students selectively focused on relevant information and performed target-oriented unidirectional transitions, undergraduate students focused on both relevant and irrelevant information and showed bidirectional transitions between various elements of the ^{1}H NMR spectra.

According to cognitive psychology, the comprehension of such external representations begins with students' perception and intake via the visual register guided by their visual attention.[11,12] Following this perception process, representational features are selected, organized, and integrated into a mental model. These cognitive processes are not arbitrary, as students must simultaneously select which features are most relevant for a particular context (i.e., selection), build a meaningful mental model (i.e., organization), and use prior knowledge to make sense of the representation (i.e., integration).[13] By doing so, students connect visual features in a representation to corresponding concepts using *sense-making* and *perceptual fluency*.[14]

For instance, to determine the R/S configuration of a chiral carbon atom, a student must use structural features (i.e., the attached substituents to the chiral carbon), connect them to chemical conventions (i.e., assigning priorities based on atomic

Received: June 24, 2025
Revised: October 3, 2025
Accepted: November 14, 2025

Table 1. List of Calculated and Incorporated Metrics

Metric	Description	Metric inclusion
Total fixation count[a]	*Total fixation count* refers to the total number of fixations within an AOI and can be calculated by summing up all fixations within an AOI. It reflects how frequently a student processes information.[61]	Excluded[b]
Total fixation duration[a]	*Total fixation duration* refers to the total amount of time a student's gaze remains within an AOI and can be calculated by summing up the duration of all fixations within an AOI. It reflects the time information is processed.[47]	Included
AOI hits[a]	*AOI hits* are the total number of times a student's gaze enters an AOI, regardless of how long or how often they fixate on it. It reflects how frequently a student refers to specific information.[62]	Excluded[b]
Transitions[a]	*Transitions* are the number of one-way gaze shifts from one AOI to another. They reflect the scanpath or sequential information processing order.[48,63]	Excluded[b]
Pattern frequency[a]	*Pattern frequency* refers to the number of occurrences for every possible three-AOI sequence reflecting specific sequential events within the relevant AOIs.[8] A *pattern frequency* is calculated by summing up every event of transitions between three AOIs that reflect a specific visual behavior. For determining the R/S configuration of a chiral carbon, relevant patterns are *switching between the chiral carbon and substituents* and *switching between substituents*.	Included
Transition entropy	*Transition entropy* reflects the randomness of *transitions* with lower values indicating biased, structured, and less random *transitions*.[64,65] The *transition entropy* was calculated in R with the GrpString Package,[66] using the Shannon entropy formula.[67]	Included
Transition–fixation ratio	The *transition–fixation ratio* reflects the degree of distribution across AOIs with lower values indicating more focused behavior and is calculated with $\textit{transition - fixation ratio} = \frac{\textit{length of collapsed AOI sequence} - 1}{\textit{length of uncollapsed AOI sequence}}$. This metric is a modified version of the fixation−transition ratio,[4] allowing for the analysis of eye movements with AOI sequences with a length of 1. To investigate structural sites locally, consecutive AOIs in the uncollapsed AOI sequence that represent the rest of the molecule (see Figure 4, gray AOIs) were merged into a single instance, allowing *transitions* to and from the rest of a molecule into account without considering irrelevant fixations.	Included
Fixation disparity	*Fixation disparity* refers to the small deviations between the gaze points of the two eyes and may reflect visuospatial cognition in organic chemistry tasks. The *fixation disparity* can be calculated by generating the covariance error ellipse area that contains 95% of the left- and right-eye difference of recorded gaze points within relevant AOIs.[41]	Included

[a]Additionally, relative values were calculated by dividing each subdimension (e.g., *total fixation count* on a single AOI) by the total sum of all subdimensions (e.g., the *total fixation count* across all AOIs).
[b]Excluded due to correlation or multicollinearity with other metrics.

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

numbers), and establish the absolute configuration (i.e., determining directionality with the Cahn−Ingold−Prelog priority rules[15]). Therefore, a student's ability to interpret a representation depends on perceiving its features and characteristics, applying cognitive processes to reason with it, drawing on conceptual understanding as well as prior knowledge, and the interplay between all these factors.[16,17]

Yet, developing these abilities can be challenging, as chemistry representations model 'invisible' entities and processes. Numerous studies have indicated that organic chemistry students often face challenges when using structural representations. They either struggle to derive implicit properties from the explicit representation, as their focus is guided by single or salient surface features, or familiar patterns, or—often related—they apply intuitive judgments, memorization, misconceptions, or rule-based reasoning.[18−39] For example, Braun and Graulich[33] reported that students rely more on the superficial indication of lone pairs when deciding on resonance stabilization than on implicit properties. While some of these strategies may lead to successful problem-solving, students' ability to attend to relevant visual features of a representation and connect it with the underlying, implicit information needs to be supported for more productive problem-solving.[14]

To support students in the adequate use of representations, it is crucial not only to understand how they are using representations and which problem-solving behaviors characterize productive use but also to what extent their visual behavior change (or do not change, i.e., remains consistent) in response to learning. Again, eye-tracking offers a powerful means of investigating such changes, as it allows capturing the degree of changes in students' eye movements and, in turn, shifts in their information processing stimulated by instructions.[40−42] For example, in a study by Hansen et al.,[40] students had to critique the chemically relevant features of accurate and inaccurate animations as well as static images of redox and precipitation reactions. After the intervention with varied animations and eye-tracking feedback, students showed increased focus (i.e., higher *relative total fixation count*) on relevant representational features.

While these studies have provided valuable insights, they often focus on single eye movement aspects (e.g., *fixation duration*) that describe distinct characteristics of students' information processing. Choosing fixation duration may be apparent when differences in this metric between students are sufficient to draw a conclusion about their problem-solving behavior. However, learning may induce eye movement changes that cannot be fully understood by isolated aspects, e.g., by choosing one metric alone. Given that eye-tracking data is inherently rich in spatial, temporal, and psychophysiological information, a more integrated analysis incorporating multiple eye-tracking metrics could offer more comprehensive access to how students develop expertise with representations in chemistry. Such an integrated analysis may support longitudinal observations of eye movement changes for evaluating instructional settings or the effectiveness of eye-tracking-based adaptive feedback systems. Accordingly, this article contributes a methodological perspective by proposing an integrative approach for analyzing eye movement changes, aiming to capture the multifaceted nature of changes in students' information processing with chemical representations across two different settings.

CAPTURING INFORMATION PROCESSING OF REPRESENTATIONS VIA EYE-TRACKING

As the brain is unable to simultaneously process all features of a representation, visual information is filtered by *visual attention*.[43] The allocation of *visual attention* can be driven by three, often interrelated, processes: 1) stimulus-based (bottom-up) processes where the attention is guided by salient features outside of the observer's control, 2) goal-based (top-down) processes where the attention is directed intentionally or volitionally by the observer's objective or task, or 3) history-based processes where the attention is guided by previous experiences of the observer (e.g., associations in which the selection of a target was rewarded in the past).[12,44,45]

While *visual attention* modulates visual signals in the information processing of a representation, representational competencies modulate how students interpret, connect, generate, select, and use representations.[46] Rau[14] describes two broad, interrelated representational competencies involved in students' conceptual processing of representations—*sense-making* and *perceptual fluency*. Whereas *sense-making* refers to the process of deliberately connecting representations to concepts and establishing meaningful connections across representations, *perceptual fluency* entails similar processing but in a quick and effortless way.[14]

Taken together, to solve organic chemistry problems, students rely on visual (e.g., the allocation of *visual attention* across a representation) and conceptual (e.g., the interpretation of a representation via *sense-making*) information processing to identify relevant features for a particular context.

This process of information processing can be captured using eye-tracking methodology.[3,47] An eye-tracker records which features a participant looked at, for how long, and in what order within a visual stimulus. These data can be used to calculate various informative metrics to draw conclusions about information processing.[48,49] Metrics such as *fixation duration*, *fixation count*, and *Area of Interest (AOI) hits* are particularly useful at providing insights into whether specific features of the representation are being processed.[47] In contrast, metrics such as *pattern frequency*, *transitions*, *transition−fixation ratio*, and *transition entropy* may reveal underlying problem-solving processes, as these metrics reflect the distribution and sequential relationships of eye movements across features of the representation.[48] Additionally, a metric such as *fixation disparity* may serve as an indicator of cognition since it captures subtle variations in eye movements, possibly indicating cognitive effort during problem-solving (for a description of those metrics, see Table 1).[41]

In the last decades, these eye-tracking metrics have been used to characterize numerous aspects of chemistry problem-solving processes, such as students' strategies, use of representations, or cognitive demand.[7−10,21,50−56] For example, Cullipher and Sevian[8] investigated how students' underlying assumptions guide their interpretation of infrared spectra. They linked the distribution of specific *pattern frequencies* to the assumptions that students made. Students with advanced conceptual sophistication focused mainly on the relevant features of the spectra, such as axes and peaks. In contrast, students with lower conceptual sophistication transitioned more frequently between molecular features of the two given structures or returned to the task question several times. Further, Hinze et al.[9] investigated low and high prior knowledge students' use of more familiar ball-and-stick models vs less familiar electrostatic potential maps to

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

solve problems related to electrostatic properties and chemical interactions. Students' *mean relative fixation duration* showed that in both contexts, high prior knowledge students used more likely the electrostatic potential maps in comparison to low prior knowledge students. Additionally, the researchers reported that even though students initially relied on the more familiar ball-and-stick models, students with higher prior knowledge began to draw on the novel electrostatic potential maps, which is reflected in their change in eye movements.

These studies exemplify how eye-tracking research can be used to capture differences in eye movements of students with varying expertise levels,[4−10] attempting to characterize developing expertise in information processing. However, comparing eye movements of students with different expertise levels only indirectly represents the development of expertise as those eye movements reflect interindividual rather than intraindividual differences. Only a few chemistry education research studies have analyzed these differences intraindividually, capturing changes in eye movements stimulated by instructions on the same students.[40−42] For example, Koh et al.[41] compared students' *fixation disparity* while solving mental rotation and pattern comparison chirality tasks before and after playing Mahjong for a month to investigate how periodically engaging in pattern-matching activities influences students' visuospatial cognition. Similarly, Moen et al.[42] compared students' *saccade count* and *saccade amplitude* while solving mental rotation tasks involving ball-and-stick models to investigate how mental rotation training with 3D block cube arrangements influences students' encoding of the ball-and-stick representations.

Although these studies contribute valuable insights, they focus on single eye-tracking metrics that capture specific aspects of students' information processing. However, developing expertise may lead to varying and multifaceted eye movement changes (i.e., changes toward more productive eye movement behavior) that cannot be fully tracked by isolated metrics. Given the inherently rich spatial and temporal nature of eye-tracking data (e.g., *fixation count* conveys spatial information about eye movements, while *pattern frequency* conveys sequential information), a more comprehensive approach that integrates multiple eye-tracking metrics, may yield more holistic access to students' developing expertise with representations.

■ RESEARCH QUESTIONS

This methodological investigation aims to explore how changes in eye movements can be characterized through an integrative approach that considers their multifaceted nature. To this end, we draw on two separate studies—one conducted without instruction and one with—to illustrate how a more integrative approach can help document the extent of change in students' eye movements. The following research questions guided our investigation:

RQ1: How can an aggregated measure that integrates multiple eye-tracking metrics be used to characterize changes in students' eye movements in settings with and without instruction?

RQ2: How do single aspects of students' eye movements (i.e., individual eye-tracking metrics) contribute to a comprehensive understanding of students' eye movement changes?

■ RESEARCH DESIGN

Experimental Setup

Two exploratory studies were conducted to characterize changes in students' eye movements across two tasks on the R/S configuration of chiral carbons, using an integrative analysis of eye-tracking data. Participants in both studies were organic chemistry beginner students (organic chemistry 1). At this early stage of expertise development, students may exhibit greater variability in eye movements and may be more responsive to instructions, making them particularly suitable for examining how the extent of changes in eye movements can be characterized in educational settings.

In study I, students solved two tasks in a direct series to explore the extent of eye movement changes without being instructed in between. In study II, students received an eye-gaze-augmented retrospective (i.e., a retrospective cued by eye-gaze replays and accompanied by reflection prompts (described by Langner and Graulich[57])) as an instruction between the tasks to explore the extent of eye movement changes comparatively (see Figure 1). For a detailed description of each study, see the

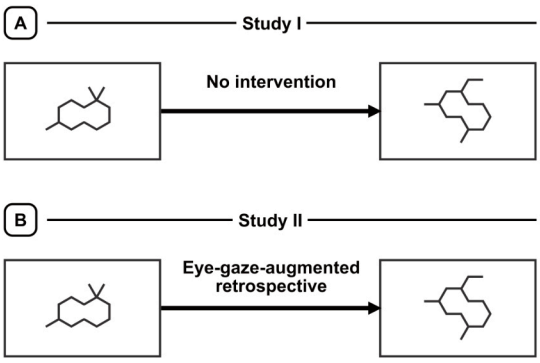


Figure 1. Two exploratory studies were conducted. Study I conducted without an instruction (A) and study II with an an eye-gaze-augmented retrospective (described by Langner and Graulich[57]) as an instruction between the tasks.

sections Study I: Changes in Students' Eye Movements without Instruction and Study II: Changes in Students' Eye Movements with Instruction.

In both studies, students were given the same tasks in which they had to determine the R/S configuration of the chiral carbons. Such visual and cognitive tasks provide a particularly suitable context to analyze developing expertise with representation through changes in eye movements. Visually, a chiral center can be processed in multiple ways, requiring careful examination of the atomic priorities and spatial arrangements. Unlike other tasks that may rely on familiar pattern recognition when interpreting representations, like mechanism problems or ranking tasks,[18,23,24] determining R/S configuration needs intentional and structured visual behavior. Students must sequentially scan specific molecular features to identify the chiral carbon atom, assign substituent priorities, and determine their directionality. Conceptually, students can employ different strategies to determine the R/S configuration. The rules provided by Cahn, Ingold, and Prelog,[15] which are taught in the students' courses can be used, but alternative methods, such as an even number of group interchanges[58] or the ±1,2,5 method[59] also exist. Each of these strategies demands different visuospatial reasoning. With this, determining the R/S configuration entails a high visual and cognitive effort, making it suitable to capture eye movement changes.

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

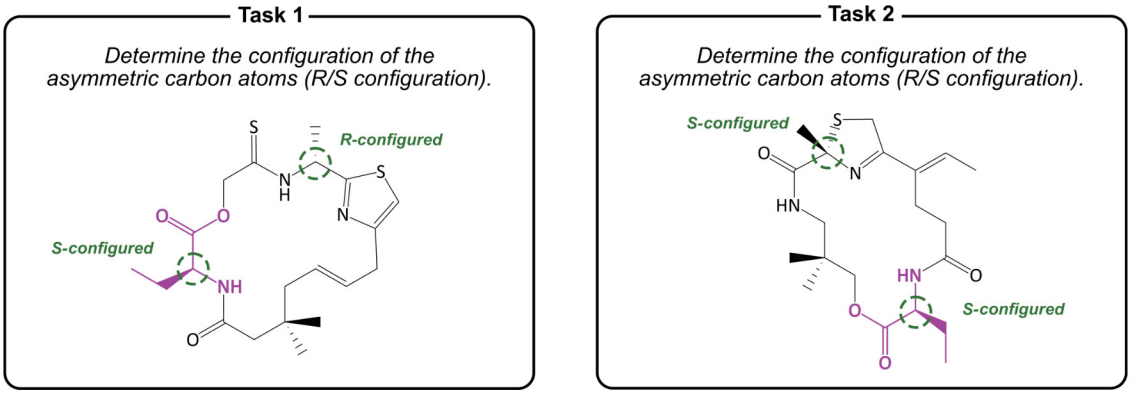


Figure 2. Design of the chirality tasks. The task solution is depicted with green circles. The chiral carbon atoms and their substituents, which are structurally identical in both tasks, are highlighted in purple. To compare participants' changes in eye movements between the tasks, only this identical structural site (purple) was selected for further analysis. The task solution, as highlighted by the green circle, was not shown to the students.

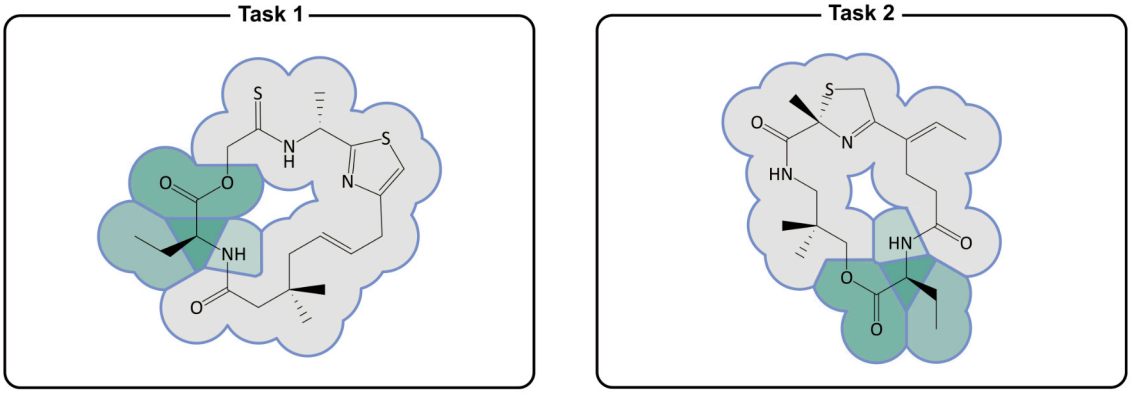


Figure 3. AOIs were defined with the limited-radius Voronoi tessellation with the radius set to the bond length. AOIs were merged to correspond to the chiral carbon and each substituent (highlighted in teal), while the rest of the molecule was combined into a single, larger AOI (highlighted in gray).

Task Design

In both studies, students were given the same two tasks. The prompt in each task was to *Determine the configuration of the asymmetric carbon atoms (R/S configuration)*. The structural representations of molecules were large cyclic organic molecules resembling macrolides, a class of natural products (Figure 2). The usage of these designed molecules, created specifically for the purpose of both studies, had several advantages. First, the representations offer a certain realism due to their similarity to the natural product group of macrolides. Second, the large cyclic structure reduces the probability of gaze guidance (e.g., visually processing the representation in the reading direction), ensuring students engage more in visual search behavior. Most importantly, the designed molecules allowed us to incorporate identical structural sites into the task set. This ensured that, on a local level, the tasks were identical, enabling a comparison of students' eye movements at this comparable structural site, while on a global level, the tasks remained distinct, minimizing the risk of a practice effect.

To ensure task clarity and that the task difficulty was appropriate for the prior knowledge of the students, we discussed the appropriateness of the task design with one of the Organic Chemistry 1 course instructors and conducted a pilot study with a small sample beforehand.

CHARACTERIZING CHANGES IN EYE MOVEMENTS

Eye-Tracking Data Collection

While students completed the tasks, their eye movements were tracked. The data collection was employed with a Tobii Pro X3-120 screen-based eye-tracking system (120 Hz sampling rate) with a 24-inch presentation monitor, a Tobii Pro External Processing Unit, and the associated software Tobii Pro Lab in a live viewer setup. The participants were instructed to sit at a distance of 60 to 65 cm from the presentation monitor and to minimize head and body movements during each task to reduce data loss. The calibration of the system was carried out with a nine-point calibration and a four-point validation. Immediately after the calibration, the gaze of the students was tracked. First, the task prompt was presented to the students. Once the task prompt was understood, the students continued by pressing a key. Then, the students saw a fixation cross for 2 s before the molecule was automatically presented. The presented molecule was scaled so that the distances between the atoms corresponded to an eye movement of approximately 2°,

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

exceeding the accuracy of the eye-tracking system of 0.7°, as well as the recommended 1−1.5° by Holmqvist et al.[48] Hence, it can be assumed that fixations can be allocated to the viewed atoms. After the task was completed, the students ended it by pressing a key again.

Data Preparation

To compare participants' changes in eye movements across the tasks, only the structural site, which was identical in both tasks, was selected for further analysis (see Figure 2). This local approach allowed for a comparison of visual behavior while minimizing the influence of differences in the rest of the stimuli. To ensure eye-tracking data quality, recordings of both tasks with eye-tracking data loss exceeding 20% were excluded from the analysis.

Participants' accuracy of responses at the selected structural site was scored. Students received a score of 2 if they correctly identified the chiral center and accurately determined the configuration, a score of 1 if they correctly identified the chiral center but inaccurately determined the configuration, and a score of 0 if they did not identify the chiral center.

Eye-Tracking Data Preparation

The AOIs were defined with the limited-radius Voronoi tessellation method with a radius set to the bond length of the molecules—a method that constructs AOIs that represent the closest area to an atom.[60] Fixations located in an AOI can be assigned to the corresponding atom since these fixations fall in the area closest to this atom within the radius of a bond length (for a detailed discussion on this method, see Langner et al.[60]). The AOIs within the identical structural sites were merged to correspond to their structural features, while the rest of the molecule was combined into a single larger AOI (Figure 3).

Calculating Eye-Tracking Metrics

To characterize visual behavior through an integrated approach and to capture the richness of spatial, temporal, and psychophysiological information in the eye-tracking data, multiple metrics for participants' eye movements were incorporated into this analysis. Table 1 lists all of the metrics along with their descriptions.

However, since some of these measures are interrelated and convey similar information (e.g., *fixation count* and *fixation duration*[68,69]), correlation analysis and multicollinearity analysis of unidimensional metrics (metrics that represent a single measure and do not include multiple measures, e.g., *transition entropy*) or subdimensions of multidimensional metrics (metrics that represent a specific measure within a multimeasure metric, e.g., *total fixation duration* on one of the AOIs) were conducted on those related metrics to exclude redundant metrics, preventing the disproportionate influence of similar information in further analysis (see Table 1). Task 1 data was chosen for correlation analysis and multicollinearity analysis, as it serves as a point of origin. For correlation analysis, Spearman correlation was chosen, as Shapiro−Wilk tests for normality revealed that most metrics were not normally distributed ($p < 0.05$). Metrics were excluded if pairs of unidimensional metrics, or the majority of pairs of subdimensions of multidimensional metrics, had a Spearman correlation coefficient exceeding 0.7 or falling below −0.7, with $p < 0.05$, indicating a statistically significant strong positive or negative correlation.[70] For multicollinearity analysis, the Variance Inflation Factor analysis was chosen. For this, a regression model was constructed using one metric of interest as the dependent variable and the other metrics of interest as predictors. Subsequently, the Variance Inflation Factor was then calculated from R^2. Metrics were excluded if they had a Variance Inflation Factor exceeding 10, suggesting strong multicollinearity among metrics.[71] Without excluding those metrics, the subsequent analysis would be skewed by a disproportionate influence of specific eye movement aspects (i.e., overemphasis). For example, in our case, if total fixation count—which correlates with total fixation duration—would not have been excluded in our analysis, students who changed their total fixation duration with a bigger extent than students who did not would be considered with a higher change in eye movements since both total fixation duration and total fixation count are taken into account to characterize the change of eye movements, even though they convey similar information. Hence, this step is mandatory to ensure a balanced influence of metrics to characterize the changes in eye movements. The same metrics were excluded for both studies. The final metrics for the subsequent analysis are indicated in Table 1.

Change in Students' Eye Movements

To determine the overall change in students' eye movements through an aggregated measure (research question 1), the average absolute bounded normalized change across all eye-tracking metrics between the tasks Δ_{Gaze} was calculated for each participant individually. First, the absolute bounded normalized change of each subdimension of the metrics, normalized by total magnitude, $\Delta_{metric} = \left|\frac{metric_{task2} - metric_{task1}}{metric_{task2} + metric_{task1}}\right|$ was calculated. The result is a unitless absolute normalized value between 0 and 1. This allows for direct comparisons across changes in eye-tracking metrics regardless of their range, unit, or direction. Additionally, using a bounded value ensures that increases and decreases in magnitude are treated symmetrically. Moreover, it helps avoid division issues when either value is zero. However, if both values are zero, Δ_{metric} is set to zero. Second, to ensure that unidimensional and multidimensional metrics are weighted equally when calculating Δ_{Gaze}, the average absolute bounded normalized change of the subdimensions of the metrics was calculated as $\overline{\Delta_{metric}} = \frac{1}{N}\sum_{i=1}^{N}\Delta_{metric,i}$. Finally, the average absolute bounded normalized change across all eye-tracking metrics was calculated as $\Delta_{Gaze} = \frac{1}{N}\sum_{i=1}^{N}\overline{\Delta_{metric,i}}$. For a practical guide on calculating Δ_{Gaze}, please refer to the Supporting Information.

Conceptually, Δ_{Gaze} is a unitless measure of relative difference within an individual's eye movements bounded between 0 and 1. A value of 0 indicates a perfect replication of gaze behavior across tasks, meaning no measurable change—an outcome unlikely given the natural variability of eye-gaze behavior. Values approaching 0 reflect high similarity, while values approaching 1 reflect increasingly different gaze patterns. The theoretical maximum of 1 occurs only in extreme cases, such as when gaze activity is entirely absent in one task and present in the other. It is important to note that Δ_{Gaze} is inherently dependent on the specific set of metrics included in its calculation. Consequently, direct comparisons of Δ_{Gaze} values across studies are meaningful only when the same metrics and computational procedures are used. With this, Δ_{Gaze} as a composite measure of relative difference does not describe initial eye movements and does not indicate the extent to which individual eye-tracking metrics change. Hence, Δ_{Gaze} offers the opportunity to compare students' eye movement change interindividually, regardless of their individual initial and end eye movements, which is useful

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

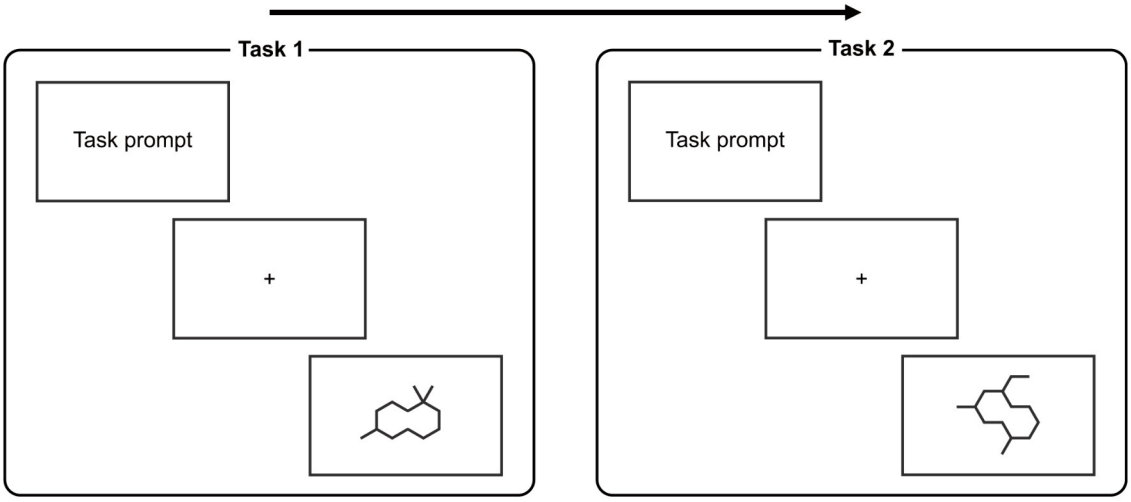


Figure 4. Study design of study I, where students solved two tasks in immediate succession. Slides of the tasks were presented in the following order: task prompt, fixation cross, and task molecule.

for settings where the individuality of students' visual processing is taken into account.

However, since Δ_{Gaze} is an aggregated measure and does not describe specific aspects of students' eye movement changes in detail, we also qualitatively compared changes in individual eye-tracking metrics (research question 2).

■ STUDY I: CHANGES IN STUDENTS' EYE MOVEMENTS WITHOUT INSTRUCTION

Participants and Setting

Students were recruited voluntarily from two Organic Chemistry 1 courses at a university in the southeastern United States in 2023. The courses, taught by different instructors, followed a traditional functional group-focused curriculum, with Klein's Organic Chemistry[72] serving as the course textbook. Both courses covered the content relevant to this study (i.e., the assignment of R/S configurations using the Cahn–Ingold–Prelog priority rules[15]).

As a compensation for their participation, students received extra points toward their final exam grades. The students had an option to earn bonus points either by participating in this study or by completing an alternative assignment. The recruitment process, data collection, and data analysis procedures followed the ethical standards approved by the Institutional Review Board (IRB) (IRB-FY24-113). All participants were informed about their rights, the option to opt out at any time, and the handling of their data. Informed written consent was obtained prior to data collection.

A total of 53 students participated in the study. After data exclusions (see Data Preparation section), 38 students were included in the final analysis. In the cohort, 34 students self-identified as female and four students as male. The average age of the participants was 21.9 ± 5.0 years, and all of the participants had normal or corrected-to-normal vision. All students were native English speakers. The data collection was conducted in English.

Data were collected in single eye-tracking sessions as part of a larger study. Students solved the two tasks in immediate succession (Figure 4). The participants were asked to solve the tasks mentally before recording their answers on a sheet of paper that depicted the respective molecule. This approach was chosen to ensure, first, minimal head and body movements during data collection and, second, that concurrent verbalizations of responses did not influence students' cognitive load or visual behavior.[61]

Results and Discussion

In study I, the overall change in students' eye movements, described by the average absolute bounded normalized change across all included eye-tracking metrics Δ_{Gaze} across all students, was 0.32 ± 0.13 (see Figure 5). While a value of 0.32 might suggest a low-to-moderate magnitude of change, Δ_{Gaze} is bounded between 0 and 1, and no established threshold yet exists to define what constitutes a small or large value in this context. Accordingly, Δ_{Gaze} should currently be understood as a comparative rather than absolute measure, gaining interpretative value only when contrasted across data sets, for example, between students, tasks, or studies within the same context.

Students' eye movement changes ranged from 0.09 to 0.60. Since the identical chiral carbon atoms are differently embedded in the molecules shown in the tasks (see Figure 2), we assumed that changes in students' eye movements are, to some extent, driven by the stimulus-based attention allocation.[12] However, we did not expect the extent of variability observed in these changes. Given that specific eye movements are linked to success,[7,21,51,54,73−75] it is highly likely that a change in success, e.g., from unproductive to productive problem solving, results in greater changes in students' eye movements.

Interestingly, this relation of change in success and increased change in eye-movement did not emerge between students' score change and Δ_{Gaze}. In study I, most of the students did not change their success scores and retained their initial score in task 1 as well as in task 2 (Figure 5). That students in study I maintained their scores is not surprising, since students did not receive instruction in between tasks. One could assume that they would have a low change in Δ_{Gaze} even though they were mostly unsuccessful in both tasks. Nevertheless, they showed quite a range of Δ_{Gaze}. This could be interpreted as an indicator that

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

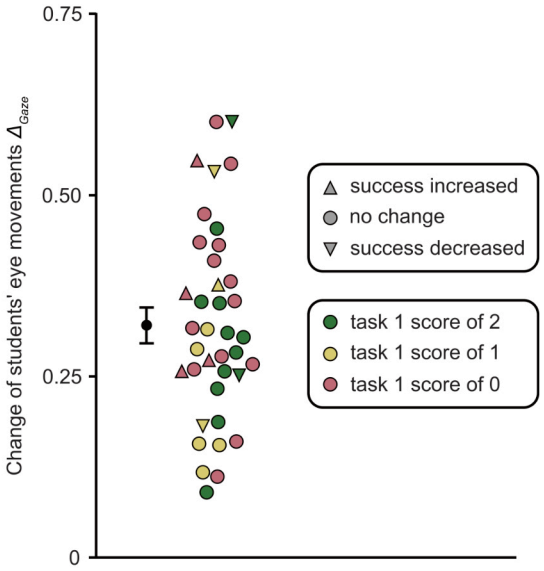


Figure 5. Overall change of students' eye movements in study I described by the average absolute bounded normalized change. Increases and decreases in success (a gain or loss of one or two points) are aggregated.

these students are in a stage of developing expertise, i.e., still figuring out how to approach the identification of the R/S configuration. Students' use of representations, while still learning, is highly individual and depends on their developing *sense-making* and *fluency* skills,[14] visuospatial abilities,[76,77] as well as their previous experiences.[12,44,45] For example, some consistently successful students may be certain in their problem-solving approach and are more guided by top-down processes, applying their strategies consistently and showing smaller changes in eye movements. Others, though equally successful, show higher changes, indicating that they may not yet have an established or stable strategy in how they typically approach such problems and may be more influenced by bottom-up processes, leading to less consistent use of specific strategies or rules. Similarly, the magnitude of changes in eye movements made among consistently unsuccessful students varied. Aligned with observations of Tóthová et al.[56] and Tóthová and Rusek,[78] the task demands in these cases likely exceeded students' capabilities, and therefore, their eye movements may reflect random search rather than consistently deliberate information processing, as their strategy repertoire is limited. These observations suggest that the magnitude of eye movement changes is not necessarily directly tied to deliberate information processing or problem-solving success.

This finding highlights the importance of considering multiple measures when evaluating the development of expertise or cognitive change, as eye movements may reveal shifts that performance metrics alone cannot detect, rather than relying on single metrics that distinguish between low and high performers.

Since Δ_{Gaze} is an aggregated measure, a more careful examination of the involved individual metrics is necessary to provide more detailed insights. To illustrate the changes in students' eye movements in detail, we present the cases of Danielle and Benjamin (Figure 6). Like most participants (30 out of 38 students), both maintained their scores across the tasks, with Danielle scoring 2 and Benjamin scoring 0 in both tasks.

Even though Danielle distributed her visual attention more evenly across the substituents (i.e., a more balanced *fixation duration* across substituents AOIs) and demonstrated more visuospatial cognition (i.e., higher *fixation disparity*), overall, Danielle showed a similar visual pattern across both tasks (i.e., small changes in *pattern frequency of switching between the chiral carbon and substituents*, *pattern frequency of switching between substituents*, *transition–fixation ratio*, and *transition entropy*) (see Figure 6A). Considering all of these aspects, her eye movement changes suggest that Danielle made only minor changes to her visual strategy across the tasks, maintaining her initial successful approach of continuously shifting between the central carbon and its residues, relating them to each other. Her changes in eye movements, described by Δ_{Gaze}, mainly result from changes in *fixation disparity*.

In contrast to Danielle, Benjamin started focusing more on the central carbon and the ethyl substituent (i.e., a higher *fixation duration* on these AOIs). He was making more consecutive transitions between the carbon and substituents (i.e., a higher *pattern frequency of switching between the chiral carbon and substituents* and a higher *transition–fixation ratio*) in the post-task. In contrast, he did not exhibit direct comparisons between the substituents (i.e., lower *pattern frequency of switching between substituents*) (see Figure 6B). Taking all of these factors into account, it can be concluded that Benjamin adjusted his visual behavior to anchor his gaze more frequently on the chiral carbon center, repeatedly shifting between the central carbon and its residues, comparable to the strategy employed by Danielle.

Both cases reflect observed tendencies in the sample, substantiating that variability in eye movements is not necessarily tied to performance outcomes but may signal different stages of developing expertise. Danielle is an example of a student with a relatively stable visual strategy, consistent with a developing or consolidating problem-solving approach. In contrast, Benjamin represents students with a less stable visual strategy, which could reflect attempts to establish a problem-solving approach that has not yet led to task success.

Danielle and Benjamin are exemplary students who did not undergo any intervention between the tasks. Therefore, their changes in eye movements are likely driven by how the chiral carbon is embedded in the molecular representation (see Figure 2), by a practice effect (e.g., task familiarity) or by variations arising from (un)established strategic approaches to solving the tasks.

The cases of Danielle and Benjamin, who share similar Δ_{Gaze} values (0.32 and 0.31), illustrate that, while an aggregated measure like Δ_{Gaze} is useful for characterizing the overall magnitude of eye movement changes between two tasks, it is essential to examine the underlying metrics in detail to draw meaningful conclusions about how those eye movements actually changed.

■ STUDY II: CHANGES IN STUDENTS' EYE MOVEMENTS WITH INSTRUCTION

Participants and Setting

Students were recruited voluntarily from an Organic Chemistry 1 course at a German university in 2021. The course followed a traditional curriculum and was structured by functional groups, mechanism types, and structure–property-relationships, with Buddrus and Schmidt[79] serving as the course textbook. The

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

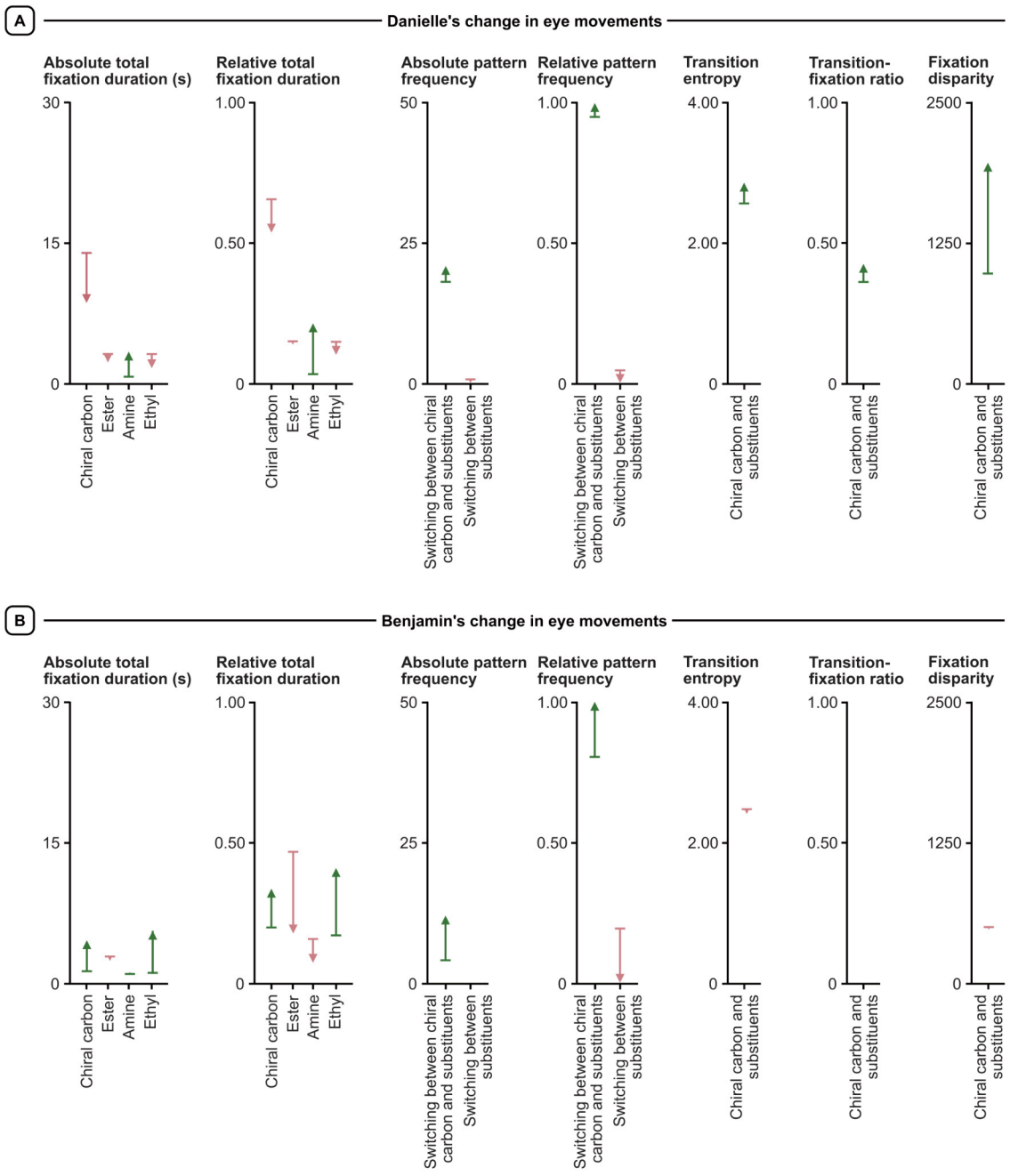


Figure 6. Changes in single eye movements metrics of consistently successful student Danielle (A) and consistently unsuccessful student Benjamin (B) from task 1 to task 2. The direction and length of the arrows indicate the changes in each individual metric.

course covered the content relevant to this study (i.e., the assignment of R/S configurations using the Cahn–Ingold–Prelog priority rules[15]).

Students from the German university did not receive any form of compensation for their participation in the study. While IRB approval is not a requirement at German universities, the recruitment process, data collection, and data analysis procedures followed ethical standards. All participants were informed about their rights, the option to opt out at any time, and the handling of their data. Informed written consent was obtained prior to the data collection.

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

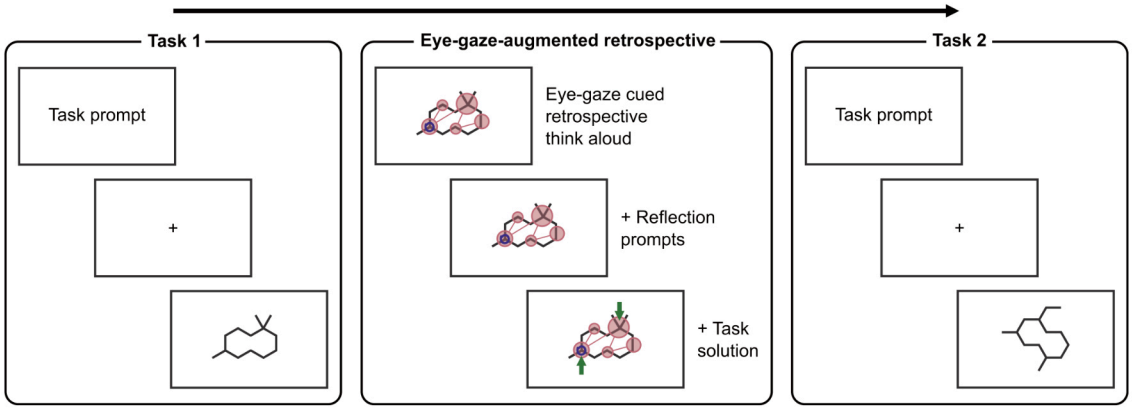


Figure 7. Study design of study II, in which students solved two tasks with the eye-gaze-augmented retrospective described by Langner and Graulich[57] as an intervention between them. The task slides were the same as those used in study I and were presented in the same order.

In total, 22 students participated in the study. Following data exclusion (see Data Preparation section), 18 students remained in the final analysis. From these students, 14 students self-identified as female and four students as male. The average age of the participants was 24.0 ± 3.6 years, and all had normal or corrected-to-normal vision. All of the students were native German speakers. The data collection was conducted in German.

Data were collected in single eye-tracking and interview sessions as part of a larger study. The task set that was carried out consisted of the same two tasks that were used in study I (see Figure 2), but with an intervention between them. During the tasks, the participants' eye movements were tracked with the same procedure as in study I. As an intervention, an eye-gaze-augmented retrospective was carried out (see Figure 7). During the eye-gaze-augmented retrospective, the participants reflected on their problem-solving process on different structural sites of the molecule of task 1 while watching their own eye-gaze replay. The eye-gaze-augmented retrospective consisted of three consecutive phases: 1) an eye-gaze-cued retrospective think-aloud, followed by 2) reflection prompts, and finally, 3) the addition of the task solution along with further reflection prompts.

In a pilot study, students demonstrated that they could interpret fixations and saccades depicted in the eye-gaze replay after a brief explanation and use the replay to reflect on their problem-solving process.

This study was previously reported by Langner and Graulich,[57] with a focus on students' reflections during the eye-gaze-augmented retrospective. For a detailed description of the procedure and the corresponding results on students' reflection, see Langner and Graulich.[57]

Results and Discussion

In study II, the overall change in students' eye movements described by Δ_{Gaze} across all students was 0.23 ± 0.09, ranging from 0.07 to 0.40 (see Figure 8). These Δ_{Gaze} values are lower than those observed in study I. A comparative discussion of both studies is provided in the Putting It All Together section. In the following section, the focus remains on the results of study II.

The majority of students maintained a score of 2 across the two tasks. In this study, consistent success appeared to be associated with smaller overall changes in the eye movements.

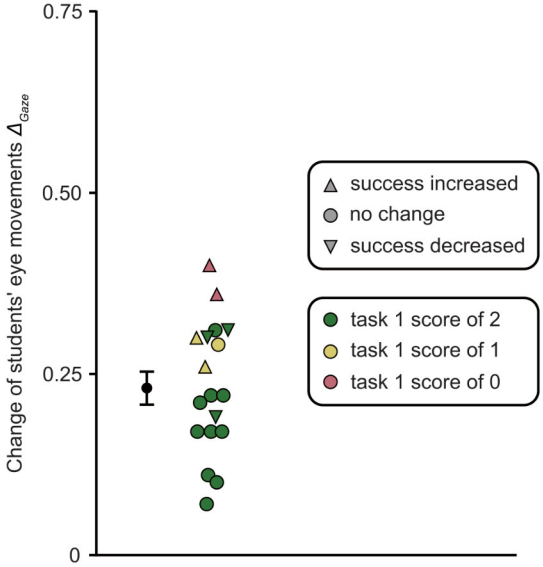


Figure 8. Overall change of students' eye movements in study II described by the average absolute bounded normalized change. Increases and decreases in success (a gain or loss of one or two points) are aggregated.

The intervention seemed to encourage previously successful students to maintain or reinforce consistent visual behavior (i.e., lower Δ_{Gaze}) across tasks, while stimulating previously unsuccessful students to adjust their visual behavior to a greater extent (i.e., greater Δ_{Gaze}) (see Figure 8). This observation aligns with the qualitative findings of this study as reported by Langner and Graulich,[57] who observed that, during the intervention, successful students often reported that they did not perceive the need to change their behavior. In contrast, the intervention mainly stimulated unsuccessful students to primarily reflect on their visual behavior, leading them to formulate intentions to focus more on relevant structural features, look more closely, or proceed more sequentially.

This suggests that the intervention stimulated successful students to reproduce or even reinforce productive visual

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

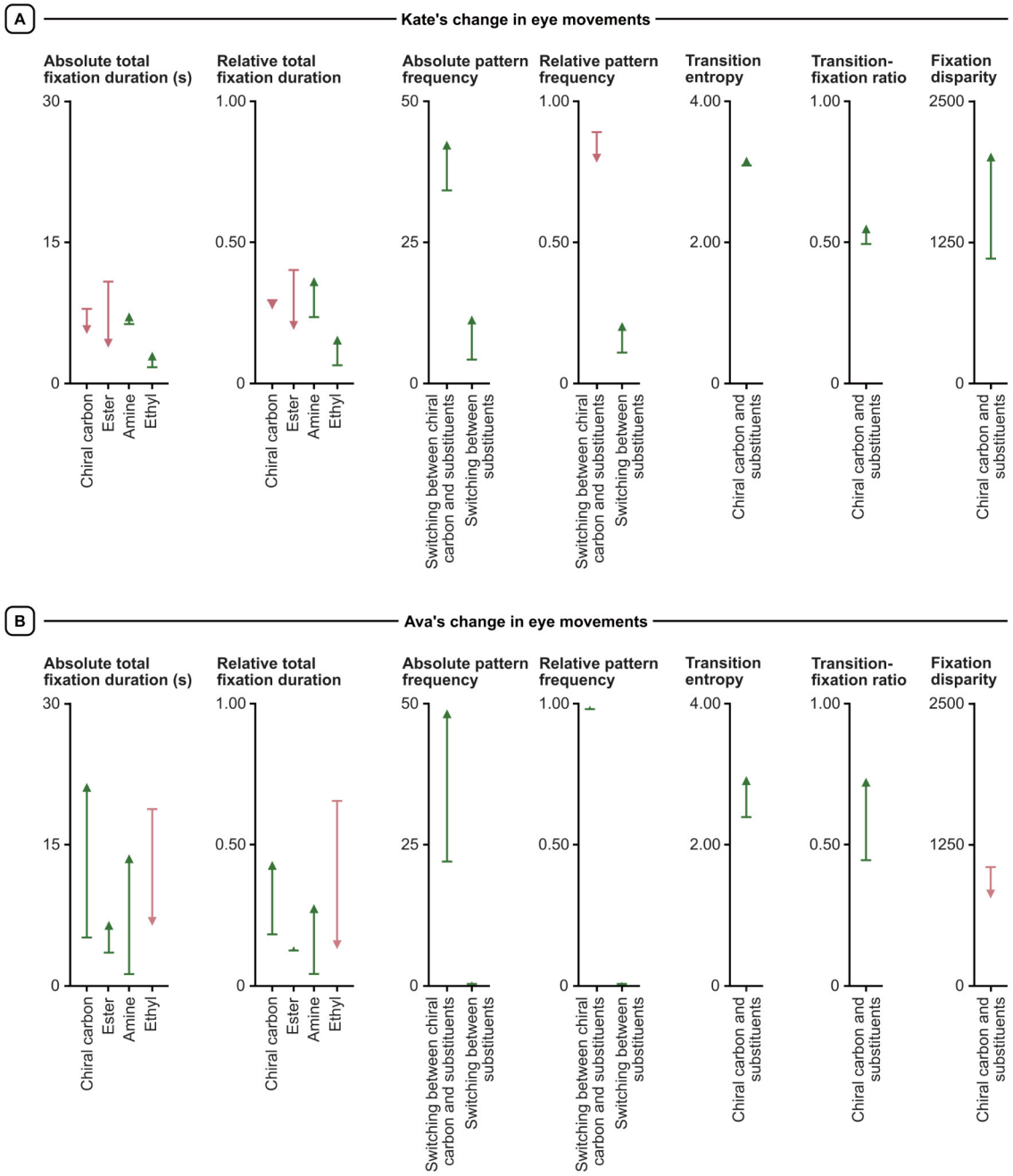


Figure 9. Changes in single eye movement metrics of Kate (A) and Ava (B) from task 1 to task 2. While Kate remained successful across both tasks, Ava improved from being unsuccessful in task 1 to successful in task 2. The direction and length of the arrows indicate the change in each individual metric.

behavior, while unsuccessful students were prompted to change their visual behavior more strongly—as seen by the cases of Kate and Ava (see Figure 9). Kate, with a Δ_{Gaze} of 0.21, was consistently successful across both tasks, while Ava, with a Δ_{Gaze} of 0.40, improved from being unsuccessful in task 1 to successful in task 2.

Kate distributed her *visual attention* more evenly across the chiral carbon and its substituents (i.e., a more balanced *fixation duration* across AOIs), while making more consecutive transitions between the carbon and substituents (i.e., a higher *pattern frequency of switching between chiral carbon and substituents* and *pattern frequency of switching between sub-*

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

stituents), without altering randomness and distribution of her gaze (i.e., small changes in *transition entropy* and *transition–fixation* ratio), and demonstrating more visuospatial cognition (i.e., higher *fixation disparity*) in the post-task (see Figure 9A). Given all of these factors, her change in eye movements suggests that she reinforced her initial strategy of continuously shifting between the chiral carbon and its substituents, as well as among the substituents themselves, to relate them to each other and determine the R/S configuration.

Unlike Kate, Ava changed her focus from the ethyl substituent to the central carbon and the amine substituent (i.e., a lower fixation duration on ethyl AOI and higher fixation durations on chiral carbon and amine AOIs). She showed more consecutive transitions between the carbon and its substituents (i.e., a higher *pattern frequency of switching between the chiral carbon and substituents* and a higher *transition–fixation ratio*), with increased randomness in her gaze (i.e., higher *transition entropy*) and less visuospatial cognition (i.e., lower *fixation disparity*) in the post-task (see Figure 9B). Based on these factors, her eye movement changes suggest that Ava changed her visual behavior to anchor her gaze more often on the chiral carbon center, thereby repeatedly shifting between the central carbon and its residues.

Taking into account that both Kate and Ava either reinforced or changed their problem-solving strategy, from an embodied and extended cognition perspective, their changes in eye movements can be interpreted as a fine-tuning of their perceptual-sensorimotor system.[80]

Both cases reflect broader tendencies in study II, where consistently successful students rather reinforced productive visual behavior, while previously unsuccessful students tended to adjust their visual behavior more strongly. Kate represents students whose lower changes in eye movements align with stable and effective visual behavior, which was further consolidated through the intervention. Ava, in contrast, is an example for students with higher changes in eye movements, whose visual behavior changed more substantially and was accompanied by improved task performance.

Study II uses the example of an eye-gaze-augmented retrospective to illustrate how Δ_{Gaze} could serve as a measure to broadly characterize the effect of an intervention by comparing students' pre- and post-task eye movements. Still, as illustrated by the cases of Kate and Ava, a clearer and more detailed understanding emerges when individual metrics are examined integratively to draw meaningful conclusions about students' behavioral change.

■ PUTTING IT ALL TOGETHER

RQ1: How Can an Aggregated Measure That Integrates Multiple Eye-Tracking Metrics Be Used to Characterize Changes in Students' Eye Movements in Settings with and without Instruction?

Both study I and study II explored how the aggregated measure Δ_{Gaze} could inform the characterization of eye movement changes, with and without an intervention in between. At first glance, the findings appear contradictory.

On the one hand, in study I, the variability of Δ_{Gaze} was notably high (see Figure 5). No clear trend emerged, as both, high and low Δ_{Gaze} values were spread across consistently successful and unsuccessful students alike—suggesting eye movement changes are not necessarily directly tied to task scores. On the other hand, in study II, students who maintained consistent performance across tasks exhibited lower Δ_{Gaze} values, indicating fewer overall changes in eye movements. Those students whose scores increased or decreased showed higher Δ_{Gaze} values, indicating more pronounced changes (see Figure 8). Still, it remains unclear whether this trend applies equally to consistently successful and unsuccessful students, as the data set primarily included consistently successful students.

Although studies I and II are not directly comparable, due to differences in the absence or presence of an intervention, students' university, course instructor, and textbook, the results of Δ_{Gaze} in both studies (see Figures 5 and 8) may serve as an initial indicator of how Δ_{Gaze} could provide access to developing expertise with representations. It is important to emphasize, however, that the studies were not intended to serve as a comparison between the groups. Rather, the aim was to characterize changes in eye movements in two distinct settings—one without and one with an intervention between the tasks. Without such an intervention, Δ_{Gaze} likely reflects a mixture of individual differences based on eye movement changes driven by a changed stimulus, a practice effect, or still developing problem-solving strategies. In contrast, an intervention like the one used in study II appears to structure students' engagement in a way that allowed Δ_{Gaze} to meaningfully reflect the extent of visual behavioral and cognitive change stimulated by the intervention.

For example, the results from both studies suggest that the intervention in study II stimulated already successful students to maintain more consistent visual behavior across the tasks, compared to the students in study I. The design of the eye-gaze-augmented retrospective and the reinforcement of accurate responses due to the task solution superimposed in the eye-gaze replay guided already successful students to allocate their attention more intentionally. This fostered a goal-driven approach while simultaneously reducing stimulus-driven attention allocation by making their visual behavior explicit and formulating intentions for subsequent problem-solving processes (for more information, see Langner and Graulich[57]).

While existing eye-tracking research already explores aggregated measures of multiple metrics—such as composite indices, machine learning-based aggregations, or regression-based modeling[81−83]—that may integratively characterize eye movements, these approaches are presently applied only to the analysis of eye movements themselves and are not yet employed for monitoring changes in them.

In contrast, other methods based on similarity algorithms—such as correlation maps, Needleman–Wunsch algorithm, or Levenshtein distance[7,84,85]—using aspects of the spatial and temporal information that eye-tracking data convey, can be used to characterizes changes in eye movements.

In Figure 10, Δ_{Gaze} is contrasted with the Levenshtein distance,[86] which was calculated from collapsed AOI sequences and normalized by the longest sequence length.[87]

While absolute mean values were higher for the normalized Levenshtein distance than for Δ_{Gaze}, both methods revealed the same relative pattern—with larger changes in study I than in study II. This convergence suggests that the observed results in both studies are robust across analytic approaches. The divergence in magnitude, however, reflects their conceptual differences. The normalized Levenshtein distance captures only specific aspects of spatial and temporal information (i.e., the sequence of AOIs visited). As a result, it is highly sensitive to scanpath variations but less sensitive to differences in other aspects such as *fixation duration*. In contrast, Δ_{Gaze} integrates

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

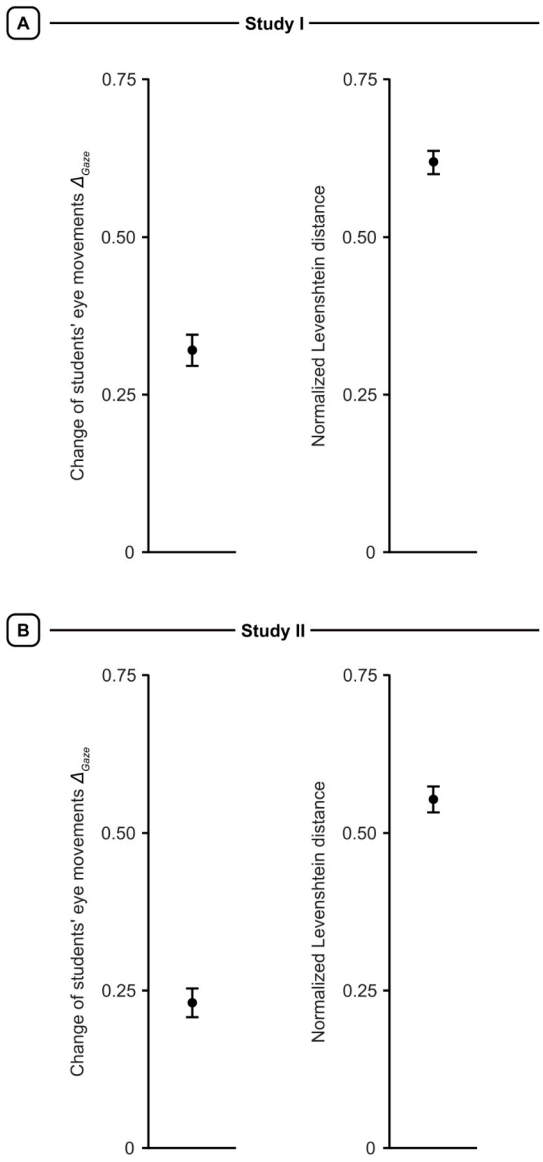


Figure 10. Mean changes in eye movements, characterized by Δ_{Gaze} (left) and by the Levenshtein distance[86] calculated from collapsed AOI sequences and normalized by the length of the longest collapsed string,[87] of study I (A) and study II (B).

multiple eye-tracking metrics, encompassing multiple spatial, temporal, and psychophysiological information, and is therefore sensitive to a broader range of variations. This does not imply that Levenshtein distance or related methods are less valuable, but rather, they are particularly advantageous in contexts where fine-grained scanpath sensitivity is required, as they can highlight nuanced changes in AOI sequences. In settings where it is not yet clear which aspects of eye movements may change due to instruction, practice, or other influences or where multiple eye movement aspects may change simultaneously, there is a lack of singular quantitative measures that encapsulate the magnitude of these changes. The Δ_{Gaze} measure addresses this gap. The use of Δ_{Gaze} can, thus, assist eye-tracking researchers in analyzing overall eye movement changes with a systematic, transparent, and easy-to-use statistical composite calculation without resorting to more complex or cumbersome methods, such as dimensionality reduction or predictive modeling.

RQ2: How Do Single Aspects of Students' Eye Movements (i.e., Individual Eye-Tracking Metrics) Contribute to a Comprehensive Understanding of Students' Eye Movement Changes?

Studies I and II demonstrated that in addition to an aggregated measure such as Δ_{Gaze}, it is essential to examine the involved metrics in detail. The cases of Danielle, Benjamin, Kate, and Ava (see Figures 6 and 9) illustrated that Δ_{Gaze} may be useful for estimating the overall magnitude of eye movement changes, but that similar Δ_{Gaze} values do not necessarily reflect the same eye movement changes. Therefore, it does not provide information about the extent of specific individual changes.

Only by considering the change of individual eye-tracking metrics together can the change in eye movements be fully characterized. This becomes evident by comparing Danielle's and Kate's changes in eye movements (see Figures 6A and 9A). Regarding their *transition entropy*, *transition–fixation ratio*, and *fixation disparity*, their changes in eye movements seem similar. However, by considering the other metrics, such as *fixation duration* and *pattern frequency*, the full picture of their eye movement changes emerges, indicating that Kate reinforced her strategy, while Danielle did not.

This detailed and comprehensive view of eye movement changes also allows for estimating the influence of instructional settings on students' eye movements. The intervention in study II modulated students' visual behavior. For example, most of the consistently successful students exhibited one or more of the following eye movement changes: 1) distributing their visual attention evenly across the chiral carbon and its substituents, as reflected in a more balanced *absolute* and *relative total fixation duration* (eight out of ten), 2) reinforcing their own productive visual strategy, as reflected in the *pattern frequency* (seven out of ten), or 3) demonstrating higher visuospatial cognition, as reflected in the *fixation disparity* (seven out of ten).

Overall, while Δ_{Gaze} provides a broad overview of changes in eye movements, a deeper and more nuanced understanding emerges when individual eye-tracking metrics are examined in conjunction. Since the choice of analyzed metrics is regularly motivated by the task type, stimulus, hypothesis, or research question,[1,62] which introduces a risk of selective reporting bias, our method mitigates this issue by incorporating multiple metrics.

Limitations

The primary limitation of our investigation is that no additional data were used to interpret students' behavior in study I. The eye-tracking data alone only allow us to infer how students processed the stimuli, but not what they were thinking.[61] While this is sufficient to illustrate the potential of our method, it does not substitute for triangulation with additional data.

Moreover, in both studies, a majority of students self-identified as female. Since participation was voluntary, this imbalance likely reflects voluntary participation rather than the overall composition of the courses. Given that gender has been associated with differences in visuospatial ability, which may, in turn, affect performance in tasks such as determining R/S configurations, this imbalance represents a limitation of our

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

studies. However, this factor was not within the scope of the present work.

Another limitation is the small sample size of study II. Since it was exploratory and required a higher degree of effort, the number of participants was kept at approximately 20 students. While this sample size might be sufficient for eye-tracking studies in chemistry education research to generate first insight to inform future research,[63] it nonetheless may be underpowered to reliably detect small- to medium-sized effects. As a result, the informative value and generalizability of the findings remain constrained. In addition, the small sample size restricts the use of inferential statistics needed to robustly examine potential differences in Δ_{Gaze} of studies I and II. Consequently, any claims about differences between those groups should be considered tentative, as they are based on holistic comparisons of descriptive quantitative data rather than on statistically supported inferences.

In addition, studies I and II are not directly comparable. Even though students in both studies took Organic Chemistry 1 courses that covered the relevant content, taught the assignment of R/S configurations using the same method, followed a traditional functional group−focused curriculum, and completed the studies in their native language with equivalent prompts, the samples in each study differ in language (English vs German), university (southeastern US university vs German university), incentive structure (extra points toward their final exam grades vs no incentive), course instructor (two US course instructors vs one German course instructor), and textbook (Klein's Organic Chemistry[72] vs Buddrus and Schmidt[79]). Hence, the differences in the studies' samples should be taken into account when comparing their result. Nevertheless, the results may serve as an initial indicator to estimate how Δ_{Gaze} could provide access to developing expertise with representations.

Furthermore, there is currently no clear threshold for what constitutes a small or high Δ_{Gaze}. Its magnitude is likely influenced by student characteristics, their performance on tasks, task properties, and the specific metrics included in its calculation. At present, Δ_{Gaze} values should therefore be interpreted within the context of the given studies or as a comparative measure rather than against a universal benchmark.

Finally, our analysis focused on changes in students' eye movements on one identical structural site in the task set. While this local approach allowed for a comparison of visual behavior, changes driven by overall differences in the stimuli limit the ability to make more general statements about problem-solving on other structural sites.

Taking these limitations into account, further research with a larger and directly comparable sample across additional chemistry domains is necessary to verify our initial observation in general and across different domains, as well as to establish normative reference ranges for Δ_{Gaze}.

■ CONCLUSION

In this exploratory investigation of two studies without and with an intervention, we demonstrated a method for how changes in students' eye movements across chemistry tasks could be characterized through an integrative approach. This is achieved by first calculating the average absolute bounded normalized change across multiple eye-tracking metrics to provide the aggregated measure of change Δ_{Gaze}. Second, the change of individual eye-tracking metrics in conjunction is considered to capture more detailed and nuanced insights.

With this, the demonstrated method considers the multifaceted nature of eye-tracking data. Δ_{Gaze} offers a comprehensive overview of overall changes in eye movement behavior and, therefore, indicates the degree of consistency in students' visual behavior when working with representations. Since this composite measure does not describe the pre and post status of eye movements and does not indicate the extent to which a single individual eye-tracking metric change, a complementary analysis of individual eye-tracking metrics is recommended.

Moreover, this systematic and transparent approach may avoid the risk of data-fishing or selective reporting bias. While prior chemistry education eye-tracking studies focused on differences in specific metrics between successful and unsuccessful students to highlight group disparities,[7,21,51,54,73−75,88] this approach risks overlooking other important eye movement aspects, such as individual adaptations in visual attention as indicators of a change in using representations or problem-solving strategies, or random variances due to a limited repertoire of strategies. Thus, by integrating multiple metrics, those aspects, which may not be directly reflected in the task score, could be captured. Considering the body of literature on eye movements and their underlying meaning in educational contexts,[3,49,89−94] a more comprehensive analysis of eye movement changes by including multiple metrics—such as the one demonstrated here—could help capture these additional aspects.

Beyond the methodological contributions, the findings also provide implications for instruction. High changes in students' visual behavior may indicate uncertainty or the absence of clear strategies, suggesting that guidance that includes visual, conceptual, and strategical cues, such as highlighting,[95,96] scaffolding,[97] eye movement modeling examples,[98] or worked examples,[99] could support students in developing more consistent and proficient use of representations. Additionally, instructions using eye-gaze augmentation and reflection may yield the potential to support students in reproducing or reinforcing productive visual behavior or changing unproductive visual behavior.

Finally, this method could support longitudinal observations of eye movement changes, enabling researchers to assess the effect of instructional settings over time. Furthermore, an integrative analysis of intraindividual learning progression using eye-tracking could contribute to the development of an adaptive feedback system based on eye-tracking—an area of increasing interest in science education eye-tracking research.[100−103]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available at https://pubs.acs.org/doi/10.1021/acs.jchemed.5c00849.

Step-by-step instructions on how to calculate the average absolute bounded normalized change across multiple eye-tracking metrics, Δ_{Gaze}, and instructions to calculate Δ_{Gaze} (PDF)

■ AUTHOR INFORMATION

Corresponding Author

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

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

Authors

Axel Langner − *Institute of Chemistry Education, Justus-Liebig-University, 35392 Giessen, Germany;* orcid.org/0000-0002-2274-8613

Marie Sahba − *Institute of Chemistry Education, Justus-Liebig-University, 35392 Giessen, Germany;* orcid.org/0009-0000-1659-5019

Maia Popova − *Department of Chemistry & Biochemistry, University of North Carolina at Greensboro, Greensboro, North Carolina 27412, United States;* orcid.org/0000-0003-0975-9534

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jchemed.5c00849

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the German Research Foundation DFG for funding this research project (project number: 453926661). We also thank the German Academic Exchange Service (DAAD), as this work was also supported by a fellowship of the DAAD. Additionally, we want to thank the students who voluntarily participated in the study, as well as Michael Schween, Pascal Klein, Larissa Hahn, and Lea Hain for their continuous support. This publication is part of the first author's doctoral (Dr rer. nat.) thesis in the Faculty of Biology and Chemistry at the Justus-Liebig-University Giessen, Germany.

■ REFERENCES

(1) Lai, M.-L.; Tsai, M.-J.; Yang, F.-Y.; Hsu, C.-Y.; Liu, T.-C.; Lee, S. W.-Y.; Lee, M.-H.; Chiou, G.-L.; Liang, J.-C.; Tsai, C.-C. A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educ. Res. Rev.* **2013**, *10*, 90−115.

(2) Tóthová, M.; Rusek, M. Eye Tracking in Science Education Research: Comprehensive Literature Review. *Sci. Educ.* **2025**.

(3) Liu, X.; Cui, Y. Eye tracking technology for examining cognitive processes in education: a systematic review. *Comput. Educ.* **2025**, *229*, 105263.

(4) Rodemer, M.; Eckhard, J.; Graulich, N.; Bernholt, S. Decoding Case Comparisons in Organic Chemistry: Eye-Tracking Students' Visual Behavior. *J. Chem. Educ.* **2020**, *97* (10), 3530−3539.

(5) Connor, M. C.; Glass, B. H.; Finkenstaedt-Quinn, S. A.; Shultz, G. V. Developing Expertise in 1H NMR Spectral Interpretation. *J. Org. Chem.* **2021**, *86* (2), 1385−1395.

(6) Topczewski, J. J.; Topczewski, A. M.; Tang, H.; Kendhammer, L. K.; Pienta, N. J. NMR spectra through the eyes of a student: eye tracking applied to NMR items. *J. Chem. Educ.* **2017**, *94* (1), 29−37.

(7) Baluyut, J. Y.; Holme, T. A. Eye tracking student strategies for solving stoichiometry problems involving particulate nature of matter diagrams. *Chem. Teach. Int.* **2019**, *1* (1), 1−13.

(8) Cullipher, S.; Sevian, H. Atoms versus bonds: how students look at spectra. *J. Chem. Educ.* **2015**, *92* (12), 1996−2005.

(9) Hinze, S. R.; Rapp, D. N.; Williamson, V. M.; Shultz, M. J.; Deslongchamps, G.; Williamson, K. C. Beyond ball-and-stick: Students' processing of novel STEM visualizations. *Learn. Instr.* **2013**, *26*, 12−21.

(10) Williamson, V. M.; Hegarty, M.; Deslongchamps, G.; Williamson, K. C., III; Shultz, M. J. Identifying student use of ball-and-stick images versus electrostatic potential map images via eye tracking. *J. Chem. Educ.* **2013**, *90* (2), 159−164.

(11) Schnotz, W. Integrated Model of Text and Picture Comprehension. In *The Cambridge Handbook of Multimedia Learning*, 2nd ed.; Mayer, R. E., Ed.; Cambridge Handbooks in Psychology, Cambridge University Press: 2014; pp 72−103.

(12) Theeuwes, J. Goal-driven, stimulus-driven, and history-driven selection. *Curr. Opin. Psychol.* **2019**, *29*, 97−101.

(13) Mayer, R. E. *Multimedia Learning*; Cambridge University Press: 2021.

(14) Rau, M. A. Conditions for the Effectiveness of Multiple Visual Representations in Enhancing STEM Learning. *Educ. Psychol. Rev.* **2017**, *29* (4), 717−761.

(15) Cahn, R. S.; Ingold, C.; Prelog, V. Specification of Molecular Chirality. *Angew. Chem., Int. Ed. Engl.* **1966**, *5* (4), 385−415.

(16) Schönborn, K. J.; Anderson, T. R. A Model of Factors Determining Students' Ability to Interpret External Representations in Biochemistry. *Int. J. Sci. Educ.* **2009**, *31* (2), 193−232.

(17) Schönborn, K. J.; Anderson, T. R. Bridging the educational research-teaching practice gap. *Biochem. Mol. Biol. Educ.* **2010**, *38* (5), 347−354.

(18) McClary, L.; Talanquer, V. Heuristic Reasoning in Chemistry: Making decisions about acid strength. *Int. J. Sci. Educ.* **2011**, *33* (10), 1433−1454.

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

(20) Rotich, F.; Ward, L.; Beck, C.; Popova, M. Attention is currency: how surface features of Lewis structures influence organic chemistry student reasoning about stability. *Chem. Educ. Res. Pract.* **2024**, *25* (4), 1071−1089.

(21) Braun, I.; Langner, A.; Graulich, N. Let's draw molecules: Students' sequential drawing processes of resonance structures in organic chemistry. *Front. Educ.* **2022**, *7*, 1−24.

(22) Braun, I.; Graulich, N. Exploring diversity: student's (un-)productive use of resonance in organic chemistry tasks through the lens of the coordination class theory. *Chem. Educ. Res. Pract.* **2024**, *25* (3), 643−671.

(23) Weinrich, M. L.; Sevian, H. Capturing students' abstraction while solving organic reaction mechanism problems across a semester. *Chem. Educ. Res. Pract.* **2017**, *18* (1), 169−190.

(24) Christian, K.; Talanquer, V. Modes of reasoning in self-initiated study groups in chemistry. *Chem. Educ. Res. Pract.* **2012**, *13* (3), 286−295.

(25) Braun, I.; Lewis, S. E.; Graulich, N. A question of pattern recognition: investigating the impact of structure variation on students' proficiency in deciding about resonance stabilization. *Chem. Educ. Res. Pract.* **2025**, *26* (1), 158−182.

(26) Domin, D. S.; Al-Masum, M.; Mensah, J. Students' categorizations of organic compounds. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 114−121.

(27) Popova, M.; Bretz, S. L. Organic chemistry students' challenges with coherence formation between reactions and reaction coordinate diagrams. *Chem. Educ. Res. Pract.* **2018**, *19* (3), 732−745.

(28) Cooper, M. M.; Corley, L. M.; Underwood, S. M. An investigation of college chemistry students' understanding of structure-property relationships. *Journal of Research in Science Teaching* **2013**, *50* (6), 699−721.

(29) 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.

(30) Graulich, N.; Bhattacharyya, G. Investigating students' similarity judgments in organic chemistry. *Chem. Educ. Res. Pract.* **2017**, *18* (4), 774−784.

(31) Graulich, N.; Hedtrich, S.; Harzenetter, R. Explicit versus implicit similarity - exploring relational conceptual understanding in organic chemistry. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 924−936.

(32) Crowder, C. J.; Raker, J. R. Patterns in Explanations of Organic Chemistry Reaction Mechanisms: A Text Analysis by Level of Explanation Sophistication. *J. Chem. Educ.* **2024**, *101* (12), 5203−5220.

(33) Braun, I.; Graulich, N. The More Explicit, the Better? How the Indication of Electron Lone Pairs Affects Students' Consideration of Resonance in Organic Chemistry. *J. Chem. Educ.* **2024**, *101* (11), 4830−4836.

(34) Galloway, K. R.; Leung, M. W.; Flynn, A. B. A Comparison of How Undergraduates, Graduate Students, and Professors Organize Organic Chemistry Reactions. *J. Chem. Educ.* **2018**, *95* (3), 355−365.
(35) Ward, L.; Rotich, F.; Raker, J. R.; Komperda, R.; Nedungadi, S.; Popova, M. Design, development, and evaluation of the organic chemistry representational competence assessment (ORCA). *Chem. Educ. Res. Pract.* **2025**, *26* (1), 244−258.
(36) 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.
(37) Graulich, N. Intuitive Judgments Govern Students' Answering Patterns in Multiple-Choice Exercises in Organic Chemistry. *J. Chem. Educ.* **2015**, *92* (2), 205−211.
(38) Bhattacharyya, G.; Bodner, G. M. "It Gets Me to the Product": How Students Propose Organic Mechanisms. *J. Chem. Educ.* **2005**, *82* (9), 1402.
(39) Rushton, G. T.; Hardy, R. C.; Gwaltney, K. P.; Lewis, S. E. Alternative conceptions of organic chemistry topics among fourth year chemistry students. *Chem. Educ. Res. Pract.* **2008**, *9* (2), 122−130.
(40) Hansen, S. J. R.; Hu, B.; Riedlova, D.; Kelly, R. M.; Akaygun, S.; Villalta-Cerdas, A. Critical consumption of chemistry visuals: eye tracking structured variation and visual feedback of redox and precipitation reactions. *Chem. Educ. Res. Pract.* **2019**, *20* (4), 837−850.
(41) Koh, D. H.; Xu, Z.; Wang, J.; Burgess, A. N.; Seccia, A.; Schneps, M.; Pomplun, M.; Lamb, R.; Keil, A.; Dawson, K.; et al. Fixation Disparity: A Possible Index of Visuospatial Cognition during Authentic Learning Tasks. *Mind Brain Educ.* **2024**, *18* (3), 346−359.
(42) Moen, K. C.; Beck, M. R.; Saltzmann, S. M.; Cowan, T. M.; Burleigh, L. M.; Butler, L. G.; Ramanujam, J.; Cohen, A. S.; Greening, S. G. Strengthening spatial reasoning: elucidating the attentional and neural mechanisms associated with mental rotation skill development. *Cogn. Research* **2020**, *5*, 1−23.
(43) Evans, K. K.; Horowitz, T. S.; Howe, P.; Pedersini, R.; Reijnen, E.; Pinto, Y.; Kuzmova, Y.; Wolfe, J. M. Visual attention. *WIREs Cogn. Sci.* **2011**, *2* (5), 503−514.
(44) Anderson, B. A. A value-driven mechanism of attentional selection. *J. Vis.* **2013**, *13* (3), 7.
(45) Awh, E.; Belopolsky, A. V.; Theeuwes, J. Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends. Cogn. Sci.* **2012**, *16* (8), 437−443.
(46) Kozma, R. B.; Russell, J. Students Becoming Chemists: Developing Representationl Competence. In *Visualization in Science Education*; Gilbert, J. K., Ed.; Springer Netherlands: 2005; pp 121−145.
(47) Just, M. A.; Carpenter, P. A. A theory of reading: From eye fixations to comprehension. *Psychol. Rev.* **1980**, *87* (4), 329−354.
(48) Holmqvist, K.; Nyström, M.; Andersson, R.; Dewhurst, R.; Jarodzka, H.; Van de Weijer, J. *Eye tracking: A comprehensive guide to methods and measures*; Oxford University Press: 2011.
(49) Skaramagkas, V.; Giannakakis, G.; Ktistakis, E.; Manousos, D.; Karatzanis, I.; Tachos, N. S.; Tripoliti, E.; Marias, K.; Fotiadis, D. I.; Tsiknakis, M. Review of Eye Tracking Metrics Involved in Emotional and Cognitive Processes. *IEEE Rev. Biomed. Eng.* **2023**, *16*, 260−277.
(50) Rodemer, M.; Karch, J.; Bernholt, S. Pupil dilation as cognitive load measure in instructional videos on complex chemical representations. *Front. Educ.* **2023**, *8*, 1−11.
(51) Karch, J. M.; García Valles, J. C.; Sevian, H. Looking into the black box: Using gaze and pupillometric data to probe how cognitive load changes with mental tasks. *J. Chem. Educ.* **2019**, *96* (5), 830−840.
(52) Tang, H.; Pienta, N. Eye-tracking study of complexity in gas law problems. *J. Chem. Educ.* **2012**, *89* (8), 988−994.
(53) Ye, J.; Zheng, Y.; Zhan, M.; Zhou, Y.; Li, L.; Chen, D. Characteristics of Pre-Service Chemistry Teachers' Mechanistic Reasoning In Organic Chemistry Tasks: An Eye-Tracking Study. *Res. Sci. Educ.* **2025**, *55*, 277−296.
(54) Tang, H.; Kirk, J.; Pienta, N. J. Investigating the effect of complexity factors in stoichiometry problems using logistic regression and eye tracking. *J. Chem. Educ.* **2014**, *91* (7), 969−975.
(55) Locatelli, S. W.; de Oliveira, L. S.; da Silva S. Junior, R.; Sato, J. R. Perspectives of the Eye-Tracking Application to Understand Access to Representational Levels in Chemistry. *J. Sci. Educ.Technol.* **2024**, *33*, 823−834.
(56) Tóthová, M.; Rusek, M.; Chytrý, V. Students' procedure when solving problem tasks based on the periodic table: An eye-tracking study. *J. Chem. Educ.* **2021**, *98* (6), 1831−1840.
(57) Langner, A.; Graulich, N. From sight to insight - reflection processes in an eye-gaze-augmented retrospective. *Int. J. Sci. Educ.* **2024**, 1−24.
(58) Idoux, J. P. A simple method for specifying the R/S configuration about a chiral center. *J. Chem. Educ.* **1982**, *59* (7), 553.
(59) Dietzel, R. A. Determination of chiral molecule configuration using the ±1,2,5 Rule. *J. Chem. Educ.* **1979**, *56* (7), 451.
(60) Langner, A.; Hain, L.; Graulich, N. Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *J. Chem. Educ.* **2025**, *102* (3), 1285−1297.
(61) Cullipher, S.; Hansen, S. J. R.; VandenPlas, J. R. Eye Tracking as a Research Tool: An Introduction. In *Eye Tracking for the Chemistry Education Researcher*; VandenPlas, J. R., Hansen, S. J. R., Cullipher, S., Eds.; ACS Symposium Series, Vol. *1292*; American Chemical Society: 2018; pp 1−9.
(62) Cullipher, S.; VandenPlas, J. R. Using Fixations To Measure Attention. In *Eye Tracking for the Chemistry Education Researcher*; VandenPlas, J. R., Hansen, S. J. R., Cullipher, S., Eds.; ACS Symposium Series, Vol. *1292*; American Chemical Society: 2018; pp 53−72.
(63) Havanki, K. L.; VandenPlas, J. R. Eye Tracking Methodology for Chemistry Education Research. In *Tools of Chemistry Education Research*; Bunce, D. M., Cole, R. S., Eds.; ACS Symposium Series, Vol. *1166*; American Chemical Society: 2014; pp 191−218.
(64) Nahlik, P.; Daubenmire, P. L. Adapting gaze-transition entropy analysis to compare participants' problem solving approaches for chemistry word problems. *Chem. Educ. Res. Pract.* **2022**, *23*, 714.
(65) Tang, H.; Day, E. L.; Atkinson, M. B.; Pienta, N. J. GrpString: An R Package for Analysis of Groups of Strings. *R Journal* **2018**, *10* (1), 359.
(66) Tang, H.; Pienta, N. GrpString: Patterns and Statistical Differences Between Two Groups of Strings; 2017. https://cran.r-project.org/package=GrpString (accessed Aug 2021).
(67) Shannon, C. E. A mathematical theory of communication. *Bell Syst. Technol. J.* **1948**, *27* (3), 379−423.
(68) Herrington, D. G.; Sweeder, R. D.; VandenPlas, J. R. Students' independent use of screencasts and simulations to construct understanding of solubility concepts. *J. Sci. Educ. Technol.* **2017**, *26* (4), 359−371.
(69) Stieff, M.; Hegarty, M.; Deslongchamps, G. Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction* **2011**, *29* (1), 123−145.
(70) Hinkle, D. E.; Wiersma, W.; Jurs, S. G. *Applied statistics for the behavioral sciences*; Houghton Mifflin: 2003.
(71) Kutner, M. H.; Nachtsheim, C. J.; Neter, J.; Li, W. *Applied linear statistical models*; McGraw-Hill: 2005.
(72) Klein, D. R. *Organic Chemistry*; Wiley: 2020.
(73) Slapničar, M.; Tompa, V.; Glažar, S. A.; Devetak, I.; Pavlin, J. Students' Achievements in Solving Authentic Tasks with 3D Dynamic Sub-Microscopic Animations About Specific States of Water and their Transition. *Acta Chim. Slov.* **2020**, *67* (3), 904−915.
(74) Slapničar, M.; Tompa, V.; Devetak, I.; Glažar, S. A.; Pavlin, J. Using an Eye-Tracker to Study Students' Attention Allocation When Solving a Context-Based Problem on the Sublimation of Water. In *Applying Bio-Measurements Methodologies in Science Education Research*; Devetak, I., Glažar, S. A., Eds.; Springer International Publishing: 2021; pp 107−127.
(75) Pavlin, J.; Slapničar, M. The Impact of Students' Educational Background, Formal Reasoning, Visualisation Abilities, and Perception of Difficulty on Eye-Tracking Measures When Solving a Context-Based Problem with Submicroscopic Representation. In *Applying Bio-Measurements Methodologies in Science Education Research*; Devetak, I.,

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

Glažar, S. A., Eds.; Springer International Publishing: 2021; pp 217−241.
(76) Hegarty, M.; Waller, D. A. Individual Differences in Spatial Abilities. In *The Cambridge Handbook of Visuospatial Thinking*; Shah, P., Miyake, A., Eds.; Cambridge Handbooks in Psychology, Cambridge University Press: 2005; pp 121−169.
(77) Kiernan, N. A.; Manches, A.; Seery, M. K. The role of visuospatial thinking in students' predictions of molecular geometry. *Chem. Educ. Res. Pract.* **2021**, *22* (3), 626−639.
(78) Tóthová, M.; Rusek, M. "Do you just have to know that?" Novice and experts' procedure when solving science problem tasks. *Front. Educ.* **2022**, *7*, 1−14.
(79) Buddrus, J.; Schmidt, B. *Grundlagen der Organischen Chemie*; De Gruyter: 2015.
(80) Pande, P. Learning and expertise with scientific external representations: an embodied and extended cognition model. *Phenomenol. Cogn. Sci.* **2021**, *20* (3), 463−482.
(81) Oubre, B.; Yang, F.; Luddy, A. C.; Manohar, R.; Soja, N. N.; Stephen, C. D.; Schmahmann, J. D.; Kulkarni, D.; White, L.; Patel, S.; et al. Eye Tracking during Passage Reading Supports Precise Oculomotor Assessment in Ataxias. *medRxiv* **2025**.
(82) Keehn, B.; Monahan, P.; Enneking, B.; Ryan, T.; Swigonski, N.; McNally Keehn, R. Eye-Tracking Biomarkers and Autism Diagnosis in Primary Care. *JAMA Netw. Open.* **2024**, *7* (5), No. e2411190.
(83) Mézière, D. C.; Yu, L.; Reichle, E. D.; von der Malsburg, T.; McArthur, G. Using Eye-Tracking Measures to Predict Reading Comprehension. *Read. Res. Q.* **2023**, *58* (3), 425−449.
(84) Tang, H.; Topczewski, J. J.; Topczewski, A. M.; Pienta, N. J. Permutation test for groups of scanpaths using normalized Levenshtein distances and application in NMR questions. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, Santa Barbara, CA, 2012.
(85) Leemans, M.; Damiano, C.; Wagemans, J. Finding the meaning in meaning maps: Quantifying the roles of semantic and non-semantic scene information in guiding visual attention. *Cognition* **2024**, *247*, 105788.
(86) Levenshtein, V. I. Binary Codes for Correcting Deletions, Insertions, and Reversals. *Sov. Phys. Dokl.* **1966**, *10* (8), 707−710.
(87) Josephson, S.; Holmes, M. E. Visual attention to repeated internet images: testing the scanpath theory on the world wide web. In *Proceedings of the 2002 symposium on Eye tracking research & applications*, New Orleans, LA, 2002.
(88) Weinrich, M.; Britt, R. Students' Attention on Curved Arrows While Evaluating the Plausibility of an Organic Mechanistic Step. In *Student Reasoning in Organic Chemistry*; Graulich, N., Shultz, G., Eds.; The Royal Society of Chemistry: 2022; pp 3−18.
(89) Alemdag, E.; Cagiltay, K. A systematic review of eye tracking research on multimedia learning. *Comput. Educ.* **2018**, *125*, 413−428.
(90) Desjarlais, M. The use of eye-gaze to understand multimedia learning. In *Eye-tracking technology applications in educational research*; Was, C., Sansosti, F., Morris, B., Eds.; IGI Global: 2017; pp 122−142.
(91) Dahlstrom-Hakki, I.; Asbell-Clarke, J.; Rowe, E. Showing Is Knowing: The Potential and Challenges of Using Neurocognitive Measures of Implicit Learning in the Classroom. *Mind Brain Educ.* **2019**, *13* (1), 30−40.
(92) Ke, F.; Liu, R.; Sokolikj, Z.; Dahlstrom-Hakki, I.; Israel, M. Using eye-tracking in education: review of empirical research and technology. *Educ. Technol. Res. Dev.* **2024**, *72*, 1383−1418.
(93) Strohmaier, A. R.; MacKay, K. J.; Obersteiner, A.; Reiss, K. M. Eye-tracking methodology in mathematics education research: A systematic literature review. *Educ. Stud. Math.* **2020**, *104* (2), 147−200.
(94) Hahn, L.; Klein, P. Eye tracking in physics education research: A systematic literature review. *Phys. Rev. Phys. Educ. Res.* **2022**, *18* (1), 013102.
(95) Schneider, S.; Beege, M.; Nebel, S.; Rey, G. D. A meta-analysis of how signaling affects learning with media. *Educ. Res. Rev.* **2018**, *23*, 1−24.
(96) Alpizar, D.; Adesope, O. O.; Wong, R. M. A meta-analysis of signaling principle in multimedia learning environments. *Educational Technology Research and Development* **2020**, *68* (5), 2095−2119.
(97) Lin, T.-C.; Hsu, Y.-S.; Lin, S.-S.; Changlai, M.-L.; Yang, K.-Y.; Lai, T.-L. A Review of Empirical Evidence on Scaffolding for Science Education. *Int. J. of Sci. and Math Educ* **2012**, *10* (2), 437−455.
(98) Xie, H.; Zhao, T.; Deng, S.; Peng, J.; Wang, F.; Zhou, Z. Using eye movement modelling examples to guide visual attention and foster cognitive performance: A meta-analysis. *J. Comput. Assist. Learn.* **2021**, *37* (4), 1194−1206.
(99) van Gog, T.; Rummel, N. Example-Based Learning: Integrating Cognitive and Social-Cognitive Research Perspectives. *Educational Psychology Review* **2010**, *22* (2), 155−174.
(100) Kennel, K.; Ishimaru, S.; Küchemann, S.; Steinert, S.; Kuhn, J.; Ruzika, S. Gaze-Based Prediction of Students' Math Difficulties - A Time Dynamic Machine Learning Approach to Enable Early Individual Assistance. *Int. J. Artif. Intell. Educ.* **2025**.
(101) Scheiter, K.; Schubert, C.; Schüler, A.; Schmidt, H.; Zimmermann, G.; Wassermann, B.; Krebs, M.-C.; Eder, T. Adaptive multimedia: Using gaze-contingent instructional guidance to provide personalized processing support. *Comput. Educ.* **2019**, *139*, 31−47.
(102) Taub, M.; Azevedo, R. How Does Prior Knowledge Influence Eye Fixations and Sequences of Cognitive and Metacognitive SRL Processes during Learning with an Intelligent Tutoring System? *Int. J. Artif. Intell. Educ.* **2019**, *29* (1), 1−28.
(103) Emerson, A.; Min, W.; Rowe, J.; Azevedo, R.; Lester, J. Multimodal Predictive Student Modeling with Multi-Task Transfer Learning. In *LAK23:13th International Learning Analytics and Knowledge Conference*, Arlington, TX, USA, 2023.

https://doi.org/10.1021/acs.jchemed.5c00849
J. Chem. Educ. XXXX, XXX, XXX−XXX

APPENDICES

Appendix A German semi-structured interview protocol of the eye-gaze-augmented retrospective in study I.

Looking back

1. *Bitte schaue dir einmal deine Blickbewegungen an. Du kannst gerne spontan deine Gedanken äußern.* [normal replay speed of the eye-gaze replay]

2. *Bitte beschreibe deinen Löseprozess anhand deiner Blickbewegungen so ausführlich wie möglich. Versuche dabei jedoch nicht die Aufgabe erneut zu lösen. Berichte dabei, was du gemacht und was du gedacht hast. Es dürfen sich Informationen doppeln. Sprich alles laut aus, egal wie irrelevant es erscheint. Solltest du unsicher über deine Erinnerungen sein, äußere dies bitte auch. Du kannst jederzeit das Video deiner Blickbewegungen pausieren, nach vorne oder hinten springen, und es dir so oft anschauen, wie du möchtest.* [replay speed was reduced to half the normal replay speed]

Awareness of essential aspects

3. *Gab es etwas das dich überrascht hat?*

4. *Gab es etwas das nicht mit deiner Erinnerung zusammengepasst hat?*

5. *Sind dir Bereiche aufgefallen, die nicht/sehr wenig/sehr stark betrachtet wurden?*

    Was könnte der Grund dafür sein?

6. *Würdest du konzeptuell oder visuell etwas anders machen, wenn du nochmal die Chance hättest diese oder eine ähnliche Aufgabe zu lösen?*

    Würdest du chemisch/inhaltlich ähnlich oder anders denken?

    Würdest du ähnlich oder anders schauen?

7. *Würdest du etwas an deiner abgegebenen Lösung verändern?*

    Was hat dich dazu bewegt deine Lösung zu verändern?

Creating alternative methods of action

8. *Gleiche deine Lösung mit der Musterlösung ab. Was schließt du daraus?* [task solution is superimposed]

    Ist die Musterlösung nachvollziehbar? Kannst du sie erklären?

    Was könnte der Grund dafür sein, dass du auf eine andere Lösung gekommen bist?

    Was könnte der Grund dafür sein, dass du diese Stelle nicht markiert hast?

    Kannst du an deinen Blicken erkennen, woran das liegen könnte?

9. *Würdest du nun konzeptuell oder visuell etwas anders machen, wenn du nochmal die Chance hättest diese oder eine ähnliche Aufgabe zu lösen?*

    Würdest du chemisch/inhaltlich ähnlich oder anders denken?

    Würdest du ähnlich oder anders schauen?

Appendix B Demographic questionnaires in the students' native language

Study I

1. Wie alt sind Sie? ☐

2. Mit welchem Geschlecht identifizieren Sie sich?

☐ weiblich ☐ männlich ☐ divers ☐ anderes Geschlecht: ☐ keine Antwort

☐

3. Beherrschen Sie Deutsch auf Muttersprachniveau?

☐ Ja ☐ Nein

4. Welches Studienfach studieren sie? (Bei Lehramt mit Fächern.) ☐

5. In welchem Fachsemeter befinden Sie sich zurzeit? ☐

6. Besuchten Sie die Vorlesung Organische Stoffchemie (OC1)? In welchem Semester zuletzt?

☐ Ja ☐ ☐ Nein

7. Haben Sie die Klausur zum Modul Organische Stoffchemie (OC1) bestanden? Wann war das?

☐ Ja ☐ ☐ Nein

Study II

1. What is your age? ☐

2. What gender do you identify as?

☐ female ☐ male ☐ diverse ☐ other: ☐ prefer not to say

☐

3. Please specify your ethnicity. (If more than one answer applies, please check more.)

☐ Caucasian	☐ African-American	☐ Latino or Hispanic
☐ Asian	☐ Native American	☐ Native Hawaiian or Pacific Islander
☐ other: ☐	☐ unkown	☐ prefer not to say

4.Do you speak English on a native level?

☐ Yes ☐ No

5. What do you major in? ☐

6. What semester are you in? ☐

7. Did you attend the lecture in Organic Chemistry (OC1)? During which semester last?

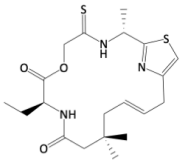
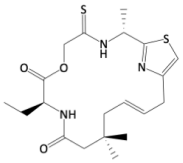
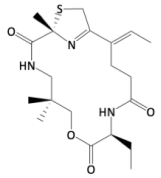
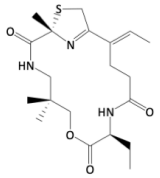
☐ Yes ☐ ☐ No

8. Did you pass the exam for the module Organic Chemistry (OC1)? What time was this?

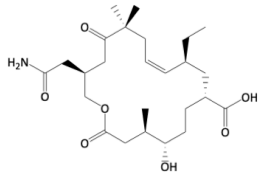
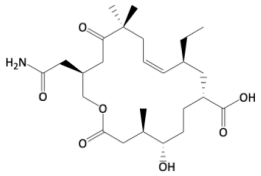
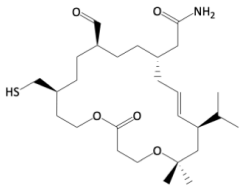
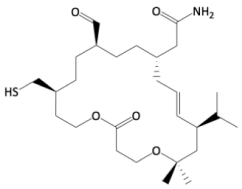
☐ Yes ☐ ☐ No

Appendix C Stimuli used in study I and II

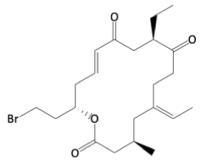
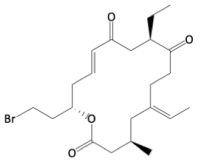
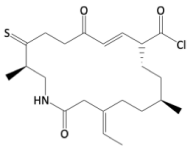
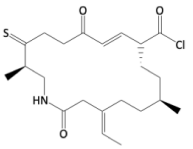
R/S configuration of asymmetric carbon atoms

	Study I	Study II
Task prompt	Bestimme die Konfiguration der asymmetrischen C-Atome (R/S-Konfiguration).	Determine the configuration of the asymmetric carbon atoms (R/S configuration).
Fixation cross	+	+
Task 1 molecule		
Task 2 molecule		

Brønsted acid-base concept

	Study I	Study II
Task prompt	Identifiziere die vier Stellen an denen das folgende Molekül am ehesten von einer starken Base deprotoniert wird.	Identify the four positions where the molecule is most likely to be deprotonated by a strong base.
Fixation cross	+	+
Task 1 molecule		
Task 2 molecule		

Electrophilicity and nucleophilicity

	Study I	Study II
Task prompt	Identifiziere die Stellen an denen sich elektrophile Zentren befinden. (Das sind solche, die von Nukleophilen angegriffen werden können.)	Identify electrophilic centers. (These are those that can be attacked by a nuleophile.)
Fixation cross	+	+
Task 1 molecule		
Task 2 molecule		

APPRECIATION

I would like to thank everyone who has supported me on the path to, and throughout, my PhD - whether professionally or personally.

Alina Majcen, Alisha Szozda, Amitabh Banerji, Anna Kira Hippert, Anna-Maria Karle, Anna Zimmermann, Annica Lau, Annika Brück-Hübner, Alexander Hopf, Alexander Maier, Anja Daßler, Annette Geuther, Argo, Benjamin Naujokat, Bettina Romberg, Brandon Yik, Büşra Tonyali, Caro Langner, Christoph Steuer, Claudia Neidel, Constantin Stremmer, Daniel Langner, Daniel Leonard Langner, David Hauck, David Hamilton, David Kranz, David Schaaf, David Weiler, Dunja Sharbat Dar, Elias (/ɪ'laɪəs/) Heinrich, Followers who constantly supported me on social media, Franz Vergöhl, Frederik Langner, Fridah Rotich, Geiler Markus, German Academic Exchange Service (DAAD), German Research Foundation (DFG), Gitte Miller Balslev, Gkiouner Mustafa, Gregor Späth, Gyde Asmussen, Heiko Barth, Hendrik Küllmar, Helena Hartmann, Helena Rist, Irina Brown, Janine Feller, Joana Konrad, Johann Seibert, John Zhou, Jule Eckhard, Julia Kliche, Julia Ortmann, Julie Szyttenholm, Justin Dillon, Justus-Liebig-University Giessen, Kathrin Haberl, Katrine Bergkvist Borch, Kerstin Kremer, Kleonie Klieber, Kookie, Larissa Hahn, Lea Hain, Lea Hotfilter, Lena Stock, LehreN-Netzwerk e.V., Leon Langner, Levi Langner, Maia Popova, Maiken Westen Holm Svendsen, Malte Michelsen, Manel Langner-Hopf, Marcel Langner, Marc Rodemer, Mareike Machleid, Marie Sahba, Marietta Nied, Martin Rusek, Marvin Rost, Melanie Seltmann, Mex, Michael Schween, Molly Atkinson, Nicolas Wöhrl, Nicole Graulich, Nicolette Macchiato, Nils Bergander, Oğuzhan Kurt, Olivia Langner, Pascal Klein, Paul Debes, Paul Reszke, Paulson Martinson, Peter Duifhuis, Petra Kretschmer, Ragi, Rebecca Bachmann, Richard Göttlich, Robbin Bernbeck, Ronja Rettich, Roshni Bhaskar, Sandra Sopay, Sascha Bernholt, Serdar Küyner, Siegfried Schindler, Silke Wagner, Silvija Markic, Simone Langner, Shanna Hilborn, Steven Petritis, Students who voluntarily participated in the studies, Sophie Freudenberg, Tamera Jones, T'ea Cameron, Timo Dahms, Timo Feick, Tina Maria Brinks, Umut Cavuldak, Valentin Engstler, Valerie Hollwedel, Zoie Bunch, and everyone I have forgotten to mention.

DECLARATION

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ in carrying out the investigations described in the dissertation.

Statement about the use of Artificial Intelligence (AI) based aids like ChatGPT or SchulKI by OpenAI, or Gemini by Google in the creation of my thesis (marked as applicable):

- ☐ I have not used any Al tool in preparing this text
- ■ I used an Al tool in the following areas (multiple answers possible):
  - ☐ Finding ideas, stimulating my creativity
  - ☐ Understanding concepts, researching facts and definitions
  - ■ Optimizing a text that I drafted myself
  - ☐ Creating entire text passages following my prompts

I used the following AI tools to improve the given passages of the text in the manner stated: ChatGPT 5.1

Date

Axel Langner

Problem-solving activities support conceptual understanding and the development of competencies needed to resolve problems and make informed decisions. In organic chemistry, problem solving relies heavily on representations that encode both explicit and implicit information. However, students often struggle to use these representations in their problem solving.

Although various instructions have been shown to support problem solving with representations, they often benefit only specific subgroups of students. In contrast, reflection on one's own problem solving—enhanced by showing students their eye movements and providing guiding prompts in an eye-gaze-augmented retrospective—offers a differentiated and personalized approach. This dissertation investigated the influence of such gaze-augmented reflection on students' problem solving in two exploratory studies—with and without the retrospective.

This investigation integrates and extends the methodological and empirical foundations established in:

> Langner A. & Graulich N. (2024). From sight to insight – reflection processes in an eye-gaze-augmented retrospective. *International Journal of Science Education*, 1–24. https://doi.org/10.1080/09500693.2024.2430804
>
> Langner A., Hain L., & Graulich N. (2025). Defining Areas of Interest in Organic Chemistry Education Eye-Tracking Research. *Journal of Chemical Education*, *102*(3), 1285–1297. https://doi.org/10.1021/acs.jchemed.4c00830
>
> Langner A., Sahba M., Popova M., & Graulich N. (2025). An Integrated Approach to Characterizing Changes in Organic. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.5c00849

The findings indicate that gaze-augmented reflection was associated with more goal-driven allocation of attention and that, despite highly individual trajectories shaped by initial problem-solving accuracy and student characteristics, accuracy converged across students. Overall, these findings highlight the potential of gaze-augmented reflection as a personalized approach to supporting students' problem solving with representations.