Chemistry Education Research and Practice

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: J. Lossjew and S. Bernholt, *Chem. Educ. Res. Pract.*, 2026, DOI: 10.1039/D6RP00153J.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/cerp

DOI: 10.1039/D6RP00153J

Tracking Student Learning Across a Conceptual Landscape: Transitions and Differential Changes in Conceptual Modes During a Unit on Chemical Kinetics and Equilibrium

Jannik Lossjew*[a] and Sascha Bernholt[a]

[a]Leibniz-Institute for Science and Mathematics Education, Olshausenstr. 62, D-24118 Kiel, Schleswig-Holstein, Germany

Abstract

Understanding chemical reactions (CR) is central to both disciplinary learning and informed engagement with societal issues in which content knowledge about CR is relevant. While prior research has mapped students' conceptual development in CR during lower secondary education, little is known about students' progression during upper secondary levels, particularly in response to targeted instruction. This study investigates how students' conceptual modes – specific reasoning patterns about the start, progress, and end of chemical reactions – develop following instruction on chemical kinetics and chemical equilibrium. Drawing on a pre-post study conducted over an extended instructional period, we assessed N = 183 German upper secondary students conceptual understanding using open-ended tasks before and after participation in a digitally implemented teaching unit. Conceptual modes were analyzed both descriptively and inferentially, incorporating difference-in-differences estimation and ordinal logistic regression to examine instructional effects and the role of prior knowledge. We included subject-related interest as an exploratory covariate. Results show significant transitions towards more sophisticated mechanistic reasoning, particularly among students who received instruction on both kinetics and equilibrium. Prior conceptual modes emerged as a strong predictor of post-instruction understanding, highlighting the cumulative nature of conceptual development as well as the diagnostic potential associated with this approach. We discuss implications for designing learning environments that scaffold mechanistic reasoning in chemistry and support the continuity of learning progressions across educational stages.

Introduction

Understanding chemical reactions is not only relevant for academic or professional pathways in chemistry but also plays a key role in preparing learners to engage with complex scientific issues such as energy transitions and environmental challenges (Anastas and Warner, 1998; Scharf and List, 2022; Schettel, 2022). In order to support learners in building up disciplinary core concepts that enable both knowledge acquisition as a basis for cumulative content learning and meaningful, contextualized application, such as engaging in informed discussions of societal relevance (Childs et al., 2015; Parchmann et al., 2006), chemical understanding needs to be developed progressively across educational stages. Approaches like learning progressions (Duncan and Hmelo-Silver, 2009) commonly suggest a spiral-curricular organization of knowledge where learners revisit core ideas with increasing depth and sophistication. One such core idea is the concept of chemical reaction (CR), which underlies key aspects of chemical reasoning and practice, including transformation, analysis, and synthesis (Sevian and Talanquer, 2014). Although cross-sectional studies have demonstrated cumulative development in students' understanding of chemical reactions over time (Boo and Watson, 2001; Emden et al., 2018; Yan and Talanquer, 2015), they also reveal persistent difficulties in developing a process-oriented understanding of chemical change – an understanding that is crucial for engaging with energetic and environmental issues in a competent way. Mastery of CR involves a highly interconnected knowledge structure and poses significant conceptual challenges, even at the secondary level (Van Driel et al., 1998). Nonetheless, such mastery is foundational for future learning in related content areas (e.g., Edelsbrunner et al., 2024). While this underscores the importance of connectable conceptual knowledge, systematic research on the development of students' understanding of chemical reactions is relatively well established at the lower secondary education level but remains sparse concerning the transition to and progression within upper secondary education. This phase of education, which bridges school and university-level chemistry, plays a crucial role in shaping learners' conceptual trajectories. Yet, to the best of our knowledge, existing studies have been predominantly cross-sectional in design and have thus failed to capture within-person development. The aim of this study is therefore to examine how upper secondary students' understanding of chemical reactions develops over time – after several weeks of instruction on chemical kinetics and chemical equilibrium, two central topics within the core concept CR. By adopting a pre-post design over an extended instructional period, we aim to trace shifts in students' conceptual modes (Yan and Talanquer, 2015) that reflect increasingly sophisticated mechanistic reasoning. In doing so, we also examine cognitive and affective predictors of this development and identify both instructional effects and persistent conceptual challenges. These findings allow us to derive implications for the development of a learning progression for upper secondary education and for designing learning environments that foster interconnected conceptual understanding of chemical reactions.

Theoretical Background

Research on chemical reactions at the lower secondary level has provided important findings, particularly through empirically validated learning progressions (LPs), which describe how students' understanding of core concepts evolves with instruction towards increasingly sophisticated ways of mastery (Corcoran et al., 2009; Duncan & Hmelo-Silver, 2009). In contrast, upper secondary education has received far less systematic attention. Based on the available evidence (Yan & Talanquer, 2005), we propose a teaching-learning sequence (TLS), understood as a research-informed instructional design approach (Bernholt & Sevian, 2018; Méheut & Psillos, 2004), which informed the instructional unit in this study.

Conceptual development of Students' Understanding of Chemical Reactions

Current State of Research. Many studies have focused on students' chemical understanding in a more global way with a strong focus on the particle model of matter (Harrison and Treagust, 2003; Krnel et al., 1998; Liu, 2001) or alternative conceptions (Garnett et al., 1995; Talanquer, 2006). However, we would like to explicitly focus on the concept CR, since comprehensive knowledge within this core concept is essential for many typical ways of thinking and working (analysis, synthesis, transformation) in our discipline. While this need for detailed understanding also applies to other areas of chemistry, chemical reactions represent an integrative element, or a *conceptual hub*, that directly connects macroscopic, particulate, and structural-energetic aspects.

While many previous studies in relation to school learning have examined specific aspects within this core concept – e.g. learning difficulties (e.g. Akkus¸ et al., 2003), misconceptions (e.g. Boujaoude, 1991; Özmen and Ayas, 2003), or specific contexts (e.g., the carbon cycle in Mohan et al., 2009) – our focus is on findings that trace conceptual development across school years of lower and upper secondary education.

In recent years, a substantial body of literature has been compiled on students' conceptual understanding of chemical reactions. For example, Hadenfeldt et al. (2014) systematically reviewed studies on students' understanding of matter, incorporating findings on chemical change. They identified hierarchically structured levels of progression, ranging from a lack of a conceptual model (e.g. García Franco and Taber, 2009) to a sophisticated understanding of the processes underlying chemical reactions (e.g. Treagust et al., 2010). Some studies have already addressed how these understandings develop over time. For instance, Boo and Watson (2001) traced how scientific explanations of reactions evolve across grade levels within the same group of learners. However, their work was limited to reactions in solution and focused on individual aspects (e.g., bond breaking, collisions, or thermodynamics) rather than their cumulative interplay. Moreover, they did not consider factors that shape students' conceptual development, such as prior knowledge or instructional strategies. More generally, many existing studies concentrate on a single point in schooling (e.g. Hesse and Anderson, 1992) or rely on cross-sectional designs (e.g. Ahtee and Varjola, 1998). Thus, Hadenfeldt et al. (2014) conclude that there is a strong need for studies that examine learning progressions.

The lower secondary level has already been well investigated (Celik, 2022; Emden et al., 2018; Walpuski and Celik, 2024), resulting in empirically validated learning progressions that suggest a gradual development of submicroscopic understanding. However, there is, to the best of our knowledge, rather limited evidence on the upper secondary level. In a comprehensive study on students' understanding of the matter concept, for example, Hadenfeldt et al. (2016) also included the concept CR. On the one hand, it was shown that across all school years that were included the understanding of chemical reactions was noticeably lower than that of the other big ideas (structure and composition, physical properties and change, and conservation). On the other hand, due to a lack of suitable learning opportunities and a lack of statistical fit, the items covering higher-level understanding (reorganization of particles and bonds, explaining and justifying reaction processes) had to be removed from the analysis. However, it is precisely the content covered by these items that elicits the difference between learning opportunities at the lower and upper secondary education level. In the latter, learners should acquire a systematic understanding of chemical reactions as the result of complex interactions of (subatomic) particles, which seems to be a learning hurdle that is difficult to overcome in the context of school learning (Ahtee and Varjola, 1998; Hesse and Anderson, 1992). This gap in research was systematically explored, at least for undergraduate students, by Yan and Talanquer (2015). The authors were able to identify overarching conceptual modes from various examples of chemical reactions in an interview study. These conceptual modes concern reasoning processes about the chemical mechanism (how a reaction takes place) as well as chemical causality (why a reaction takes place), and, at a more detailed level, focus on the start, the progress, and the end of a chemical reaction. Conceptually, a more sophisticated understanding represented by these modes is inherently tied to the topics chemical kinetics and chemical equilibrium, which we understand as relevant "progression factors" and refer to as topic areas in the following. Although Yan & Talanquer (2015) provided important insights into learners' explanatory patterns – both regarding chemical mechanism and chemical causality – the authors' findings remain predominantly qualitative. Moreover, reported results were limited to a relatively small sample that only included students who had already obtained a high school diploma (i.e., undergraduate students and beyond). It was shown that learners at all educational levels in the study tend to use agentive arguments in their explanations of chemical reactions. Encouragingly, however, the cross-sectional data also revealed developmental trends from the early undergraduate level to the advanced graduate level. While learners from the former level were generally unable to provide explanations at the molecular level, at least some learners of the latter level were able to provide mechanistic explanations including multiple interactions between various agents.

Highlighting the Research Gap. Building on the theoretical considerations outlined above, we note that there are only limited findings on students' individual conceptual development throughout secondary education. With regard to instruction, little is known about conceptual development in connection with targeted teaching within the topic areas of chemical kinetics and chemical equilibrium that we have identified. Although extensive research has addressed these factors, focusing on alternative conceptions (Banerjee, 1991; Cakmakci, 2010; Cakmakci et al., 2006; Hackling and Garnett, 1985; Hesse and Anderson, 1992; Özmen, 2008; Van Driel, 2002) and topic-specific difficulties (Akkus¸ et al., 2003; Cakmakci and Aydogdu, 2011; Justi, 2003; Kousathana and Tsaparlis, 2002; Monteiro et al., 2020; Seçken and Seyhan, 2015), the impact of systematic learning opportunities on students' sophistication in understanding chemical reactions remains widely unexplored. Research by Yan and Talanquer (2015) suggests that moving beyond the submicroscopic understanding typically

View Article Online
DOI: 10.1039/D6RP00153J

acquired by the end of lower secondary education is associated with a quantitative understanding of chemical reactions, which is particularly fostered through instruction in the two key topic areas of kinetics and equilibrium.

In the present study, we therefore examine the extent to which instruction on the identified topic areas influences the development of students' reasoning about chemical reactions. As a particular feature of our study, some classes received instruction only on chemical kinetics (**k**inetics-**o**nly **g**roup, KOG) while others also received instruction on chemical equilibrium (**k**inetics-and-**e**quilibrium **g**roup, KEG). Importantly, this differentiation was not part of the original design, as we generally recommend that students receive instruction in both topic areas, but emerged from time-related constraints in some classes. However, this "naturally" occurring variation provided an opportunity to examine the differential contribution of the two topic areas to students' developing understanding of chemical reactions. We assume differences as prior research suggest that instruction on chemical equilibrium introduces additional conceptual resources, particularly related to energetic considerations and the dynamic nature of chemical systems. Moreover, these ideas may support students in developing a more differentiated understanding of reaction systems and in distinguishing different reaction systems.

In doing so, we contribute to the existing body of research by providing insights into time-related conceptual learning within the same group of learners after several weeks of instruction (see methods). The study therefore evaluates changes in conceptual understanding following a research-informed TLS. While the empirical evidence generated at this stage of the study is not yet sufficient to claim an empirically validated learning progression, our findings can inform the design and the subsequent empirical investigation of a learning progression. In this way, our findings may help move towards a validated LP for upper secondary education, similar to those that are already existing for lower secondary education.

Shaping Conceptual Development

In addition to instructional opportunities, further cognitive and affective factors shape individual conceptual development. **Prior knowledge** in the context of chemical reactions may shape how learners interpret and benefit from subsequent learning opportunities (Etzel et al., 2025), as more advanced concepts often build on earlier forms of understanding (cumulative conceptual understanding). For instance, understanding the relationship between activation energy and the Maxwell-Boltzmann distribution presupposes an understanding of particle collisions and interactions. If these earlier ideas are not sufficiently integrated, learners may struggle to connect more advanced concepts and revert to simpler reasoning patterns. This aligns with the notion that prior knowledge is particularly influential when it is congruent with new learning content (Brod, 2021; Edelsbrunner et al., 2024). We therefore distinguish between two perspectives on prior knowledge: (1) general prior knowledge, reflected in students' previous chemistry grade, and (2) specific and congruent prior knowledge, reflected in learners' conceptual modes prior to instruction. Because conceptual modes build on one another while increasing in explanatory power, they can be understood as cumulative, suggesting that learners' understanding at the end of lower secondary education strongly shapes subsequent learning. This also highlights the importance of closely diagnosing students' explanatory patterns, particularly during transitions between school levels.

In addition, **subject-related interest** may shape conceptual development, particularly in less familiar contexts where transfer cannot be assumed (Schwartz et al., 2005). Subject-related interest has been associated with increased cognitive engagement and academic performance (Hidi & Renninger, 2006; Marsh et al., 2005). While the focus of the present studies is on cognitive learning gains in students, we included subject-related interest as an exploratory covariate to isolate instructional and prior knowledge effects from potential interest-related influences.

Research Questions and Hypotheses

Based on the theoretical considerations regarding conceptual development and targeted instruction, the following research questions were derived:

View Article Online
DOI: 10.1039/D6RP00153J

Research Question 1

To what extent do upper secondary students' conceptual modes regarding the mechanistic interpretation of chemical reactions change after working in a learning environment on chemical kinetics (and chemical equilibrium)?

Given the strong relationship between the content-related description of conceptual modes and the content taught in the context of a research-informed TLS covering these topic areas (Yan and Talanquer, 2015), we expect an observable shift in students' conceptual modes of chemical mechanism towards a higher explanatory power. However, we also expect to descriptively observe that the shift to conceptual modes with higher explanatory power is favored when learners start from a certain baseline level of understanding (Edelsbrunner et al., 2024).

Research Question 2

To what extent does this change differ for students who only received instruction on chemical kinetics (kinetic only group, KOG) and those who received instruction on both topic areas (kinetic and equilibrium group, KEG), also depending on the specific reaction example used?

As Yan and Talanquer (2015) showed, students' expressed conceptual modes are highly dependent on the chosen reaction example. We therefore also assume to observe these differences when comparing two different reaction types (complete combustion reaction and incomplete esterification) that we chose for our investigation (further details are provided in the 'instruments' section). In the case of these two reaction types, it is plausible that some learners will incorrectly assume the establishment of a dynamic equilibrium even in the case of a complete reaction. Such overgeneralized application of a newly learned concept is known as negative transfer (Woltz et al., 2000). This risk may be particularly relevant for learners in the KEG condition, as they were explicitly taught about dynamic equilibrium. However, the instructional unit explicitly categorized and contrasted different reaction systems in order to support learners in differentiating between complete reactions and equilibrium reactions. Such categorization processes can promote deeper concept formation (Goldwater et al., 2018). Therefore, despite the possibility of negative transfer, we do not expect substantial overgeneralization in the KEG group when interpreting the end of complete reactions. In addition, we expect moderate advantages in explaining the start and progress of reactions for students that have been taught both topic-related subsets (KEG), as these students had more learning opportunities to deal with chemical reactions in general, which, for example, can be interpreted as time on task (e.g. Hattie, 2010).

In turn, we do expect substantial differences between KOG and KEG regarding the interpretation of incomplete reactions. As the evaluation of systemic conditions leading to the establishment of a dynamic equilibrium, reflecting the conceptual organization from complete to incomplete chemical reactions (Van Driel et al., 1998), is only taught in KEG, we assume that there will be overall a greater shift to conceptual modes with a higher explanatory power in this group.

Research Question 3

What influence do prior conceptual modes and students' last chemistry grade have on the outcome modes when accounting for subject-related interest as an exploratory covariate?

As previously discussed, we assume students' prior conceptual modes to be strong predictors of their outcome modes because their order in explanatory power can be understood as a cumulative progression and they represent highly specific prior knowledge. This type of prior knowledge is relevant and strongly related to the new knowledge components to be acquired and should therefore have a significant positive effect on the conceptual development (Brod, 2021; Edelsbrunner et al., 2024). In addition, we also assume students' grade to have a slight to moderate positive effect on the outcome modes, since this can also be seen as a proxy for general prior knowledge. However, we hypothesize that this effect will be smaller than the effect of students' prior conceptual mode as the participants' grade is much more general and,

View Article Online
DOI: 10.1039/D6RP00153J

in this sense, less relevant and related to the new concepts. Subject-specific interest may also show a positive association with students' outcome modes and was therefore considered to control for potential confounding effects.

Methods

Ethical Considerations

Informed consent was obtained from all participants. Participation was voluntary, and confidentiality and anonymity were ensured. The study was conducted in accordance with institutional ethical standards and approved by our local ethical committee (ID 2022_33_KU).

Sample

Sample description. The sample included 14 classes from German schools, comprising 9 classes from *Gymnasiums* (academic-track secondary schools) and 5 classes from comprehensive schools. The classes were drawn from two federal states: 13 classes from Schleswig-Holstein and 1 class from Rhineland-Palatinate, thus a strong representation of northern German schools. These classes were taught by 13 different teachers, with one teacher responsible for teaching two different classes. The surveys took place during the 2022/2023 school year and covered a period from September 2022 (first pre-test) to July 2023 (last post-test). In total, the participating classes included 245 students. However, 183 students (55.1% female, 43.3% male, 1.6% non-binary) could be included in the analysis, as these were students who were present at both pre- and post-measurements. Participation was voluntary, and both school administration and students (or their legal guardians) provided informed consent for participation in the study. German was by far the most spoken language, accounting for 87.8% of the participants. This was followed by 1.1% Arabic, 1.1% Turkish, 9.5% other languages and 0.5% uncertain (e.g. participants who did not clearly specify their most spoken language). The average last chemistry grade of the sample was 2.52, which, according to the German grading system falls between "good" (2) and "satisfactory" (3). However, the full range of grades, from "very good" (1) to "insufficient" (6), was represented within the sample. Teachers were able to decide whether to participate in the study based on the topic, and due to typically non-fixed thematic sequences within upper secondary education, the grade levels (11th grade to 13th grade) varied. However, grade 11 was most frequently represented, which marks the beginning of upper secondary education in the federal states. Thus, the age range, which we did not explicitly collect, typically falls between 16 and 18 years.

Participants expected prior knowledge. Since all participants had completed the transition to the upper secondary level, we synthesized the learning-relevant competence goals that should be met by the end of lower secondary education on the basis of regional curricula from the two federal states (Ministerium für Bildung, Wissenschaft, Jugend und Kultur, 2010; Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein, 2019; Ministerium für Bildung, Wissenschaft, Weiterbildung und Kultur, 2014). A detailed summary of these goals can be found in the appendix (Section A, Table A1). Given that many of the participating students received much of their prior chemistry instruction during the COVID-19 pandemic, it is reasonable to infer that certain content-specific learning objectives may not have been fully achieved (Betthäuser et al., 2023; Lewalter et al., 2023).

Overall, both curricula introduce chemical reactions at the submicroscopic level primarily through the breaking and forming of chemical bonds. Although activation energy is included, it is usually presented superficially as an energy barrier that must be overcome, reinforcing the idea of heat as an external trigger rather than relating it to molecular energy distributions (Pölloth et al., 2023). Prior research suggests that this qualitative treatment limits students' understanding of why increasing temperature leads to more successful collisions (Justi, 2003). Despite minor differences in phrasing and exemplarity, both curricula target a similar competence level by the end of lower secondary education: students should explain reactions using the particle model, describe energy changes, and understand bond breaking and formation. Deeper mechanistic insights are reserved for upper secondary education, where kinetic and equilibrium principles are addressed more systematically. Although Rhineland-Palatinate briefly introduces first equilibrium concepts through reversibility, the dynamic

View Article Online
DOI: 10.1039/D6RP00153J

establishment of equilibrium is not treated, leaving mechanistic understanding largely underdeveloped.

Study Design

We conducted the study as pre-post design over an extended instructional period. During this period, students participated in a digitally supported, TLS-oriented teaching unit, which is described below together with the assessment opportunities forming the basis of the investigation.

Teaching Unit. The learning opportunities for the two topic areas are part of a comprehensive teaching unit that lasts several weeks (approx. 10 weeks at 2 lessons per week) and has been implemented as a digital learning environment in the learning management system (LMS) Moodle (for a structure and content overview see Table 1). In relation to the overarching goal of the project, the implementation of the digital variant was chosen to develop learning environments based on the framework by Kubsch et al. (2022). This framework provides a structured approach for integrating meaningful, inquiry-oriented learning opportunities in the sense of project-based learning and is guided by epistemic activities embedded in digital learning environments. The – in most cases novel – LMS-supported unit was piloted in spring and summer 2022 to identify implementation and comprehension challenges. Based on weekly exchanges with the participating teacher, the unit was revised and key elements for teacher PD were derived. For the main study, all teachers attended a PD session and received a detailed teaching guide with lesson plays to support implementation fidelity. Nevertheless, teacher-specific variations over the more than ten-week implementation period cannot be fully controlled.

Table 1: Structure of the digital learning environment in lesson sets with accompanying pre- and post-test and the corresponding content covered in each lesson set.

Pre-Test (approx. 90 minutes)			
Introduction to using Moodle (for learners)			
Lesson Set	**Progression Step**	**Content**	**Contexts**

View Article Online
DOI: 10.1039/D6RP00153J

1	Quantitative dynamics of chemical reactions	- Definition of the concept of reaction rate - Appropriate metrics for measuring reaction rate - Distinction between the average and instantaneous reaction rate	- Wheat flour reacts at different speeds (digesting a bread roll, rising yeast dough, flour dust explosion)
2	Control and prediction of chemical reactions	- Modelling the reaction processes using the collision model - Prediction and explanation of various factors influencing the course of reaction - Catalysis and alternative reaction pathways	- Influence of acid concentration and temperature on the dissolution of fizzy dietary supplement tablets - Food spoilage and metabolism
3	Dynamic equilibrium	- Reaction incompleteness and its causes - Systemic conditions for equilibrium adjustment - Characterization of the equilibrium state	- Production process of banana flavor
4		- Manipulation of the equilibrium state	- Industrial applications using the example of nitrogen fertilizers
Post-Test (approx. 60 minutes)			

As mentioned earlier, KOG completed lesson sets 1-2, while KEG completed all lesson sets. The four lesson sets for this unit were systematically built on the three steps of analyzing quantitative dynamics of chemical reactions, predicting and controlling chemical reactions, and analyzing dynamic equilibria to enable cumulative learning within the core concept of chemical reactions. In the instructional design, we also included reviews of conceptual difficulties (Bain and Towns, 2016; Heeg et al., 2020) and considered basic multimedia instructional strategies (Mayer, 2021) as well as specific digital tools (e.g. Lossjew and Bernholt, 2024). Table 1 shows that the content of lesson sets 2-4 was strongly informed by the empirical findings of Yan and Talanquer (2015) on how students interpret the start, progress and end of reactions mechanistically. In line with the increasing explanatory power of conceptual modes, we successively incorporated these elements into the instructional design, which therefore forms a empirically grounded TLS. Table 1 further illustrates that we embedded the domain-specific content in interrelated sub-contexts to provide relevance to the otherwise rather abstract content (Parchmann et al., 2006; Schneider et al., 2020). In summary, the unit aimed to shift students' understanding of chemical reactions from a descriptive to a predictive, process-oriented view through cumulative learning opportunities. The instructional unit can be provided by the authors as a Moodle course upon request. Translation into English or any other language than German would be required.

Available Data. The pre- and post-test were administered via the Moodle test tool. While the digital implementation allowed us to capture all learner artifacts across ten weeks, this article focuses on the pre- and post-surveys to examine conceptual development within the core concept CR following instruction in upper secondary education.

Instruments

Assessing Conceptual Modes. To assess learners' conceptual modes in the pre- and post-test, we developed five open-ended questions based on the qualitative description of different conceptual modes by Yan and Talanquer (2015) for two reaction examples (Table 2). As part of the piloting phase (see Teaching Unit), we examined the questions regarding 1) content validity, specifically whether learners addressed the intended conceptual aspects, and 2) comprehensibility. We could confirm both aspects based on the piloting results. In addition, the content validity of the items was discussed with colleagues with expertise in chemistry education. From a content perspective, we have selected a combustion reaction and an esterification reaction to cover both complete and incomplete reactions that learners are not completely unfamiliar with based on the previous curriculum. From an assessment perspective, we developed the questions to address the start of the reaction (2 questions for each reaction type), the further course or progress of the reaction (1 question each) and the end of the reaction (2 questions each), thus corresponding to the three-part classification of conceptual modes for chemical mechanism by Yan and Talanquer (2015). Based on the literature synthesis and the proposed TLS, we focused on chemical mechanism, as no substantial progression in chemical causality was expected within the scope of our instructional unit. Figure 1 shows the development of an instrument for quantitatively assessing students' conceptual modes regarding the start, progress and end of a reaction, based on an extension to the work of Yan and Talanquer (2015).

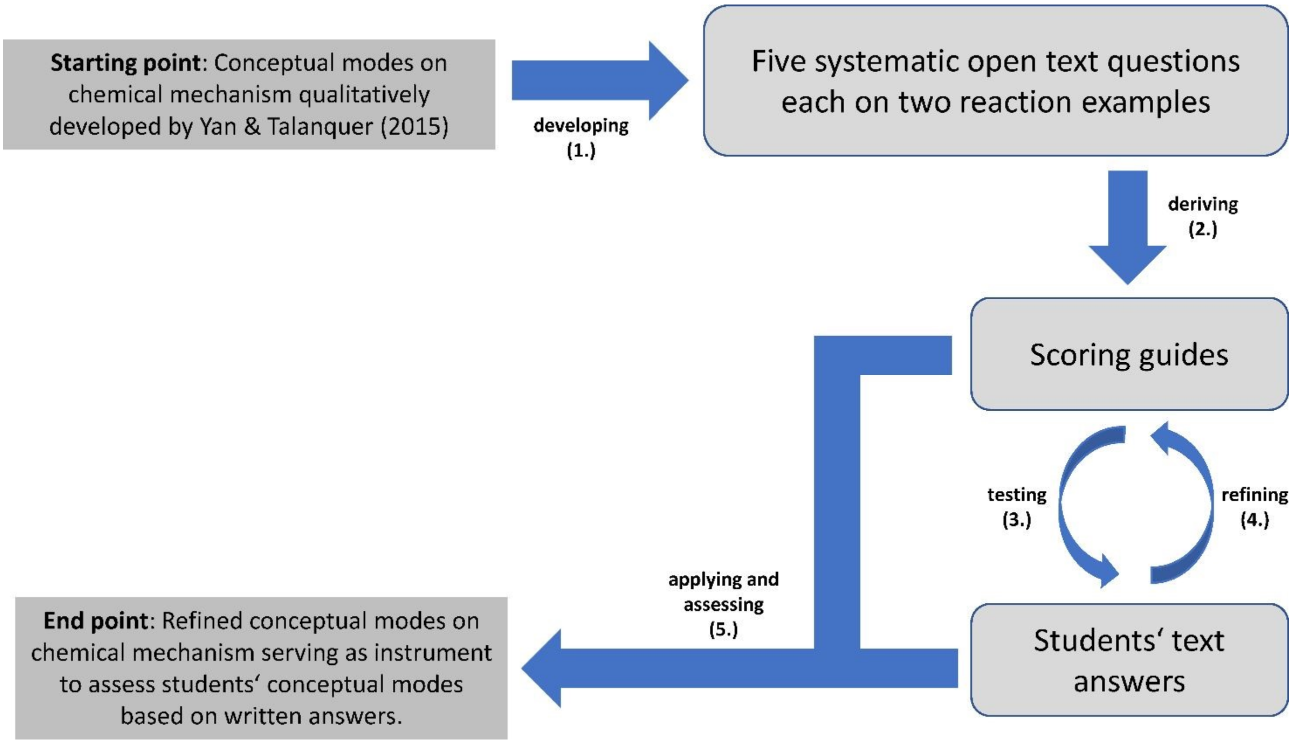

Figure 1: Process steps for designing and implementing a quantitative instrument to capture the conceptual modes on the chemical mechanism of learners in upper secondary education.

As shown in Figure 1, testing the developed scoring guides led to a refinement process. For this purpose, a substantial proportion (approximately 30%) of the free-text responses were evaluated by human raters, followed by a discussion of issues related to the differentiation of response quality. It was noted that some learners incorporated the concept of particle collisions, but in a relatively generic and undifferentiated manner. In contrast, other responses described collisions as a necessary prerequisite for further processes (e.g. bond breaking and reformation). To address this qualitative content difference, we refined the scoring guide by dividing “random collisions” and “changes through interactions” into “simple collisions” and “advanced collisions”. This refinement process results in an adapted version of conceptual modes on chemical mechanism (see Figure 1) that is based on the analysis of all written student responses.

View Article Online
DOI: 10.1039/D6RP00153J

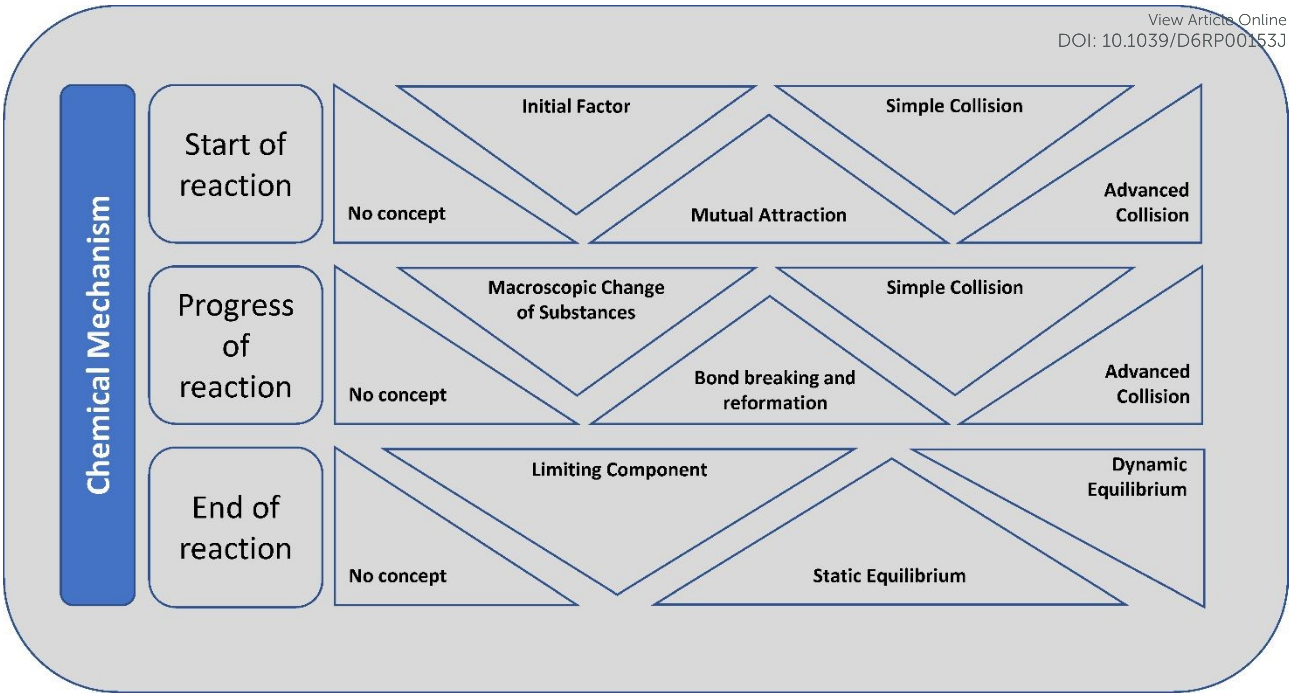

Figure 2: Refined conceptual modes on chemical mechanism, adapted from Yan & Talanquer (2015).

While two independent coders had already been involved in this refinement process as part of their training, the entire dataset was subsequently coded using the final scoring guide. Importantly, this guide did not solely rely on specific keywords but focused on the conceptual meaning of students' responses, including examples and boundary cases to account for variability in students' writing and expression. Based on sub-mode descriptions and authentic examples (see Table 2), each response was assigned to a corresponding sub-mode. Approximately 20% of the dataset was coded by both coders, and interrater reliability was calculated on this basis (see Table 2), yielding satisfactory results across all main modes. Table 2 illustrates the open-ended questions addressing the three levels of conceptual modes using the example of the combustion reaction. The detailed sub-modes are also presented, along with sub-mode descriptions and respective student examples. An analogous table for the esterification reaction is provided in the supporting information.

View Article Online
DOI: 10.1039/D6RP00153J

Table 2: Overview of modes and sub-modes with increasing explanatory power. In addition, a description, an example and the inter-rater reliability per mode are given. (*) Questions were translated from German. (**) As initial factors, we included both the mention of single compounds as an initiator and external factors. The table continues on the following page.

Mode	**Open Text Item***	**Sub-modes**	**Description**	**Student Example**	**Cohens κ**
Start (MS)	1) Describe what happens to the carbon and oxygen particles when they are heated, i.e. before the actual reaction begins, so that a reaction can occur. 2) Explain why heating the mixture causes the reaction to start immediately and why the reaction does not stop immediately when the Bunsen is removed.	Advanced Collision (**COLEXP**)	Students use the concept of random effective collisions and resulting (electronic) interactions to explain the reaction start.	Oxygen molecules move at a certain speed and collide with the vibrating carbon atoms. Once the particles randomly collide with sufficient kinetic energy (which is increased by the Bunsen), the reaction begins.	0.71 (pre); 0.81 (post)
		Simple Collision (**COLSIMP**)	Students use the concept of random collisions without further interactions.	As soon as two molecules collide, a reaction occurs.	
		Mutual Attractions (**ATTRAC**)	Students use the concept of mutual attraction to explain the start of reaction.	Attractive forces between oxygen molecules and carbon atoms cause the two particles to approach each other, initiating the reaction.	
		Initial Factor** (**INIT**)	Students use the concept of an initiating substance (reactivity) or name external factor as reason for the reaction start.	The heat from the burner breaks bonds in oxygen molecules, allowing them to react with carbon atoms.	
		No Concept (**NOCONC**)	No concepts are used to explain the reaction start.	CO_2 is formed in this reaction.	
Progress (MP)	3) Describe how you imagine the reaction at the particle level to proceed.	Advanced Collision (**COLEXP**)	Students use a combination of different interactions (continued collisions – including product particles, bond breaking/forming, escape/remaining of a substance in the system) to explain the progress.	Oxygen molecules and carbon atoms continue to collide and can rebound from each other. However, if the kinetic energy is sufficient, the collisions will lead to further reactions. Since the reaction of the particles themselves releases energy, the reaction continues and newly formed CO_2-molecules leave the open system.	0.81 (pre); 0.85 (post)
		Simple Collision (**COLSIMP**)	Students use the concept of perpetual collisions without describing further interactions.	The molecules continue to collide and thus react. They can also rebound after colliding.	
		Bond breaking and reformation (**BONDS**)	Students only use the concept of bond breaking/forming to explain the progress.	New bonds between carbon-atoms and oxygen-atoms are formed and old ones are continually broken.	
		Macroscopic change of substances (**MACRO**)	Students describe the course of the reaction from a substance-based perspective (e.g. increase in products/decrease in reactants)	More and more carbon and oxygen react with each other, so that more and more carbon dioxide is formed.	
		No concept (**NOCONC**)	No concepts are used to explain the progress.	The round flask is getting warm.	

View Article Online
DOI: 10.1039/D6RP00153J

Mode	Open Text Item	Sub-modes	Description	Student Example	Cohens κ
End (ME)	4) Describe how you imagine the end of the reaction at the particle level. Which particles are still present at this point? 5) Explain why the reaction potentially stops at this point at what factors influence the end of the reaction.	Dynamic Equilibrium (**DYNAMIC**)	Students explain that reactions do not actually end, but instead reach a state of dynamic equilibrium.	Although the ratio of carbon, oxygen and carbon dioxide no longer changes, reactions continue to take place at the particle level, in both directions at the same speed.	0.85 (pre); 0.91 (post)
		Static Equilibrium (**STATIC**)	Students explain the end of the reaction by the achievement of a static equilibrium.	At some point, the ratio of carbon, oxygen and carbon dioxide no longer changes; at this point, equilibrium has been reached and no more reactions take place.	
		Limiting Component (**LIMIT**)	Students explain the reaction end by the presence of a limiting factor and/or by a complete reaction of the reactants.	Carbon is the limiting factor, as oxygen is present in excess. The reaction stops when the limiting factor is used up.	
		No concept (**NOCONC**)	No concepts are used to explain the reaction end.	At some point, the flask is empty.	

View Article Online
DOI: 10.1039/D6RP00153J

Subject-Related Interest. To measure subject-related interest, we used a common scale for German-speaking countries, consisting of 4 items rated on a 4-point Likert scale ranging from “does not apply at all” and “fully applies” (Daniels, 2008). The scale demonstrated good internal consistency with an overall Cronbach’s α of 0.84. The item specific values ranged between 0.78 and 0.85. Example items include “I find chemistry exiting” or “I am eager to learn more about chemistry”. The analysis of internal consistency was conducted using the *psych* package in R (version 4.4.1).

Other Background Variables. In addition to the aforementioned variables, we assessed further background variables (most commonly spoken language, grade and gender), of which the last chemistry grade is used in this study as a proxy for general prior knowledge. Grades were coded with the German 1-6 grading scale, with 1 representing the best grade and 6 the worst.

Data analysis

The analyses are informed by the qualitative findings on cross-sectional development in the core concept CR by Yan and Talanquer (2015) and apply inductively derived conceptual modes in a comprehensive sample of learners before and after having received instruction on the progress-related topics chemical kinetics and chemical equilibrium. The authors elaborated the conceptual modes’ connectedness with the specific reaction example. This leads, in addition to the clear differentiation between modes (start, progress and end of the reaction) to the fact that we also did not cumulate the modes expressed by learners into an aggregated overarching ability, but have rather considered differential changes **per mode** and **per reaction** example (3x2 schemes).

Analytical Steps. To address the different research questions, we employ diverse methods that align with the specific epistemological focus of each question. Figure 3 serves as advanced organizer for the methodological triad which we will explicate in the following.

View Article Online
DOI: 10.1039/D6RP00153J

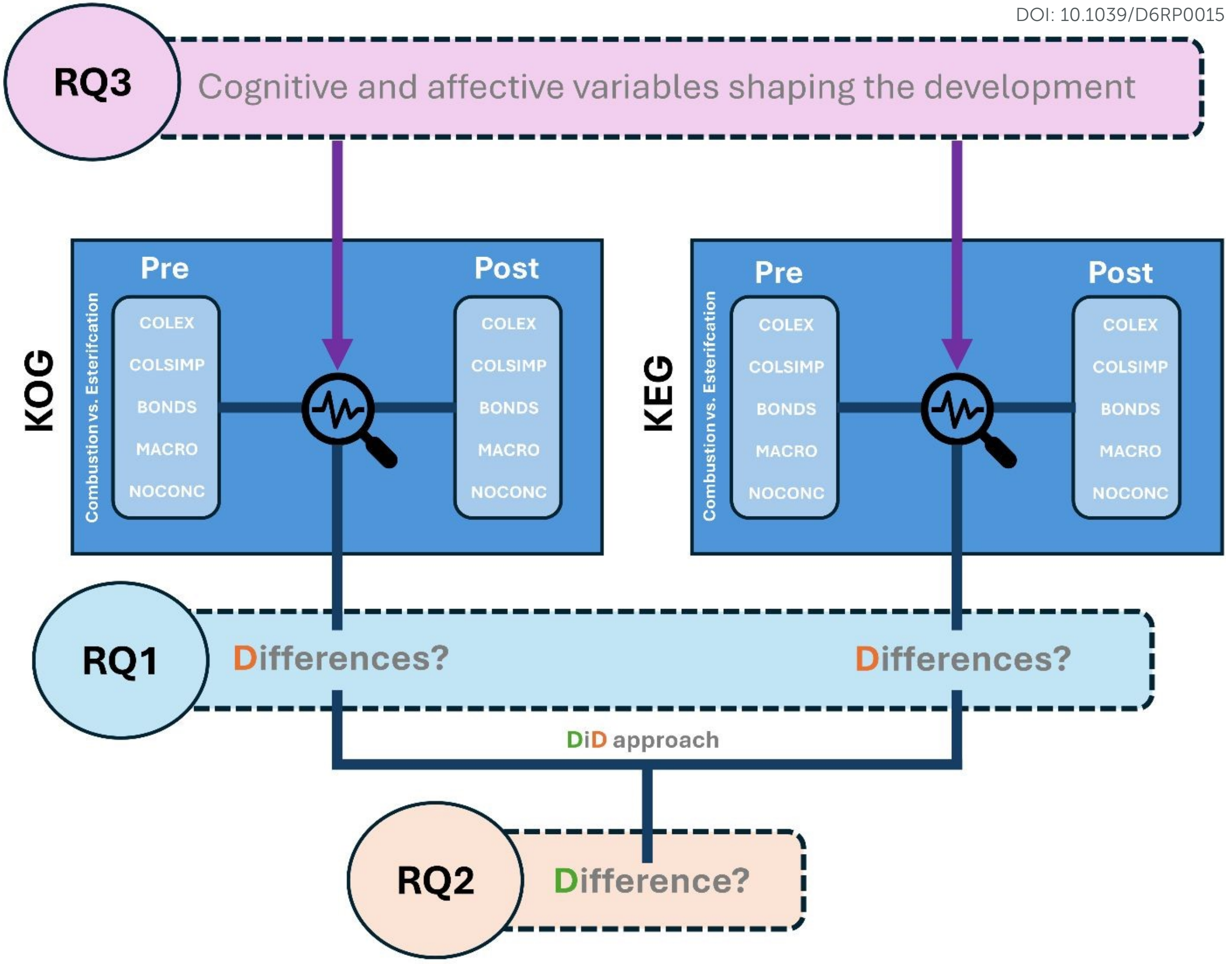

Figure 3: Overview of analytical steps linking their specific focus to the outcome measures.

First, regarding **RQ1**, we aim to analyze overall changes in the distribution of participants across the different conceptual sub-modes (cf. Table 2) before and after the teaching unit. For this purpose, we first observed shifts in students' conceptual sub-modes directly from the contingency tables. These observable shifts were identified as changes in the frequency of students in (especially higher) sub-modes as well as by the transitions between sub-modes from before to after the teaching unit. To further analyze these shifts, we calculated a test of marginal homogeneity (also known as Stuart-Maxwell test), which can be applied to paired samples to assess whether changes in the marginal distributions of squared contingency tables are statistically significant (Stuart, 1955). This analysis was performed using the package *DescTools* in R (version 4.4.1).

Second, as described in our research gap, some of the participating classes received instruction only on chemical kinetics (KOG), while another subset of students was additionally taught chemical equilibrium (KEG). This difference in instructional content effectively creates a treatment group (KEG) and a corresponding comparison group (KOG). To assess the effect of this additional instruction (**RQ2**), we analyze how the transitions between conceptual modes differ between these groups over time. Specifically, we compare the change in conceptual mode occupancies before and after instruction within each group, these represent **differences** in the conceptual modes over time. We then analyze how the magnitude of these changes differs between the two groups, i.e. group-related **differences** due to differential instruction. This approach, known as Difference-in-Differences (DiD, Graves et al., 2022) is also visualized in Figure 3 to better capture its idea. It allows us to isolate the effect of the additional instruction by accounting for general learning trends in both groups. More specifically, we employed the extended additive DID approach for categorial outcomes (Graves et al., 2022). To quantify group-related differences, an Average Effect of Treatment on the Treated (ATT) is calculated for all occupancy levels. The ATT estimates the causal effect of an intervention (in this case additional learning opportunities on chemical equilibrium) for those individuals who actually received the treatment – that is, the extent to which their probability of occupying a certain post-sub-mode is higher or lower compared to what would be expected without intervention. Further details for our choice of the additive approach and a detailed description of the ATT estimation procedure can be found in the appendix (Section C). This analysis was performed using the packages *fastDummies* and *ggaluvial* in R (version 4.4.1).

Third, we analyzed how different predictors influence the likelihood of being in a higher conceptual mode after the teaching unit (**RQ3**). While the Stuart-Maxwell test (see RQ1) can determine whether the marginal distributions before and after the teaching unit differ significantly, it does not allow for conclusions about the dependencies between pre- and post-sub-modes. However, these dependencies reflect the influence of specific prior knowledge (Edelsbrunner et al., 2024) and, thus, the potential for taking advantage of new learning opportunities (Etzel et al., 2025). Accordingly, the analysis aims to provide insights into the cumulative structure of conceptual modes. Moreover, additional control variables (chemistry grade, interest) were included to allow for the isolation of the effects of specific prior knowledge.

View Article Online
DOI: 10.1039/D6RP00153J

To integrate these considerations analytically, we conducted an ordinal regression analysis *as a final modeling step*. Ordinal logistic regression is a method used to analyze the influence of various predictors on an ordinal dependent variable – one that has a natural ranking but unknown intervals between categories. Treating conceptual modes as an ordinal outcome variable is based on the theoretical assumption that higher modes exhibit increasing explanatory power, making ordinal logistic regression the appropriate method to analyze how different predictors influence the likelihood of being in a higher mode. We developed the regression models incrementally using a manual forward stepwise logistic regression approach that was guided by theoretical considerations (see appendix for details on the model-building process). In these regression models, the pre-sub-modes (specific prior knowledge) were consistently treated as ordinal predictors. Consequently, we computed two model variants: (a) using NOCONC as the reference category to assess the predictive power of higher pre-sub-modes compared to the absence of concept activation in the pre-condition, and (b) introducing contrast coding to examine the predictive strength of each pre-sub-mode relative to the next lower category. This dual strategy allowed us to capture both the distinction from NOCONC and the incremental effects across the ordinal levels of the pre-sub-modes, thereby offering insight into their cumulative structure. In cases where the effect of a particular pre-sub-mode appeared very large, we investigated whether this was due to extremely low frequencies in the dataset. In all such instances, the respective sub-mode had very few occurrences in the pre-condition, resulting in rather inflated coefficient estimates. This issue results from the low occurrence of certain conceptual sub-modes at pre-test. Given the learners' expected prior knowledge, it is plausible that only a small number of students will have conceptually advanced modes of reasoning when they enter the instructional unit. To preserve both interpretability and robustness, we excluded such categories from the model and did not report the associated effect. However, this necessary data-driven exclusion imposes limitations on the analysis of the cumulative structure of conceptual modes. We performed several classical checks to ensure the statistical validity of our regression models, which are typical for ordinal regression (e.g. multicollinearity or checking for the proportional odds assumption). A more detailed

View Article Online
DOI: 10.1039/D6RP00153J

discussion of these aspects can be found in the appendix (Section C). Only the models that met all these criteria were retained. We controlled for instructional groups (KOG, KEG) in every final model. Since the three conceptual modes (start, progress and end of a reaction) were analyzed separately and could differ depending on the reaction example (combustion or esterification), we did not aggregate across modes or reactions examples, resulting in six separate regression analyses. All (pre-)analyses were performed using the packages *car*, *MASS*, *brant* and *Ordinal* in R (version 4.4.1).

Results

We organized this section according to the three analytical steps derived from our three research questions (Part I to Part III). Table 3 summarizes students' grades and subject-related interests across the sample. Of the total sample (*N* = 183), who participated in both the pre- and post-test, only 150 students provided responses for these background variables. As a result, the sample size for RQ3 is also reduced to *N* = 150. As mentioned before, grade was reported on a German grading scale, where 1 represents the highest score and 6 the lowest score. A comparison of these descriptive parameters of both variables in the full sample and the reduced sample is provided in the appendix (Section D) to rule out any bias caused by overachieving or overly interested learners.

Table 3: Descriptive summary of the predictors last chemistry grade and subject-related interest.

Variable	Total (N)	Mean	SD	Min	Max
Grade	150	2.52	1.15	1.00	6.00
Subject-related interest	150	2.55	0.75	1.00	4.00

Part I – Overall Changes in Students Conceptual Modes

Figure 4 shows for all main modes (start, progress and end), and for both reaction types (combustion, esterification) the pre-post distribution of the respective sub-modes, which directly corresponds to the marginal distributions (i.e., the percentage distributions) at both measurement points. Detailed transition matrices in the appendix complement these results. Overall, the increasing color intensity across all main modes and reaction examples indicates a shift toward sub-modes with higher explanatory power. While Figure 4 displays changes in

View Article Online
DOI: 10.1039/D6RP00153J

the marginal distributions (corresponding to overall changes) before and after the instructional opportunities, we also visualized the transitions between sub-modes for each main mode in Figures 5 to 7. These transitions are described in more detail below.

View Article Online
DOI: 10.1039/D6RP00153J

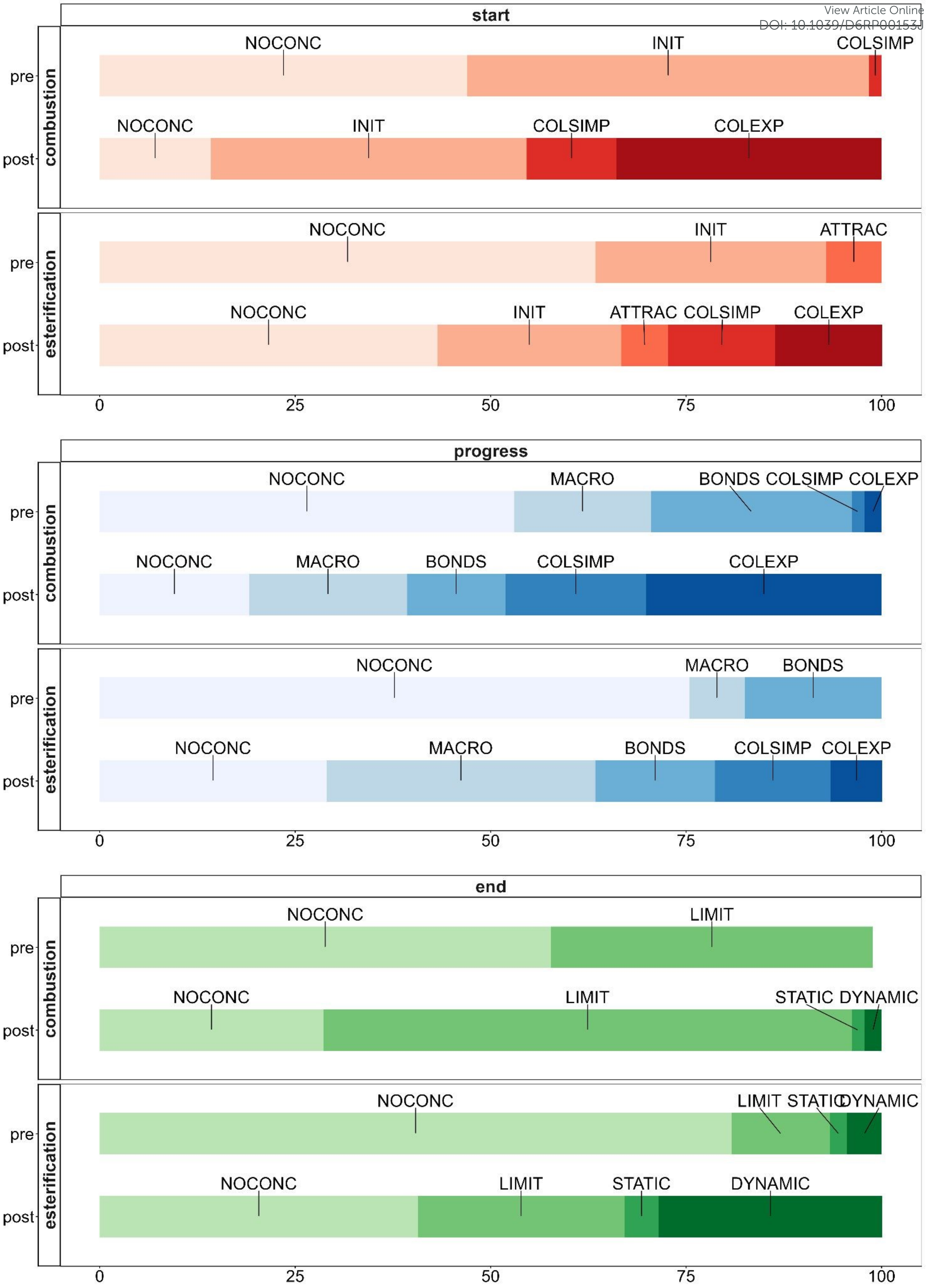

Figure 4: Pre-post distribution of sub-modes, structured by main mode (x-axis) and reaction example (y-axis) for the start (red), progress (blue) and end (green) of both reaction types.

Reaction Start. In the pre-condition, many learners do not activate specific concepts when describing the start of a reaction (red part of Figure 4). Consistent with expectations based prior knowledge (see expected prior knowledge) the idea of an initial factor (INIT) is most likely taken up in both reactions, while collision theory is largely absent. A key difference is that mutual attractions (ATTRAC) are not activated in combustion reactions but appear in esterification. Overall, esterification seems more conceptually demanding, aligning with curricular analyses that emphasize combustion reactions in lower secondary chemistry. This trend is also reflected in the sub-mode transitions from pre to post (see Figure 5). Although we can observe positive shifts for esterification overall, they are noticeably smaller compared to combustion. At the same time, Figure 5 shows that the presence of specific prior knowledge favors transitions into higher sub-modes – for example the number of transitions from INIT to a comprehensive understanding of collisions (COLEXP) compared to no concept (NOCONC) to COLEXP. However, this observation requires confirmation (see part III). Regressive transitions occur, but they are overall much less frequent.

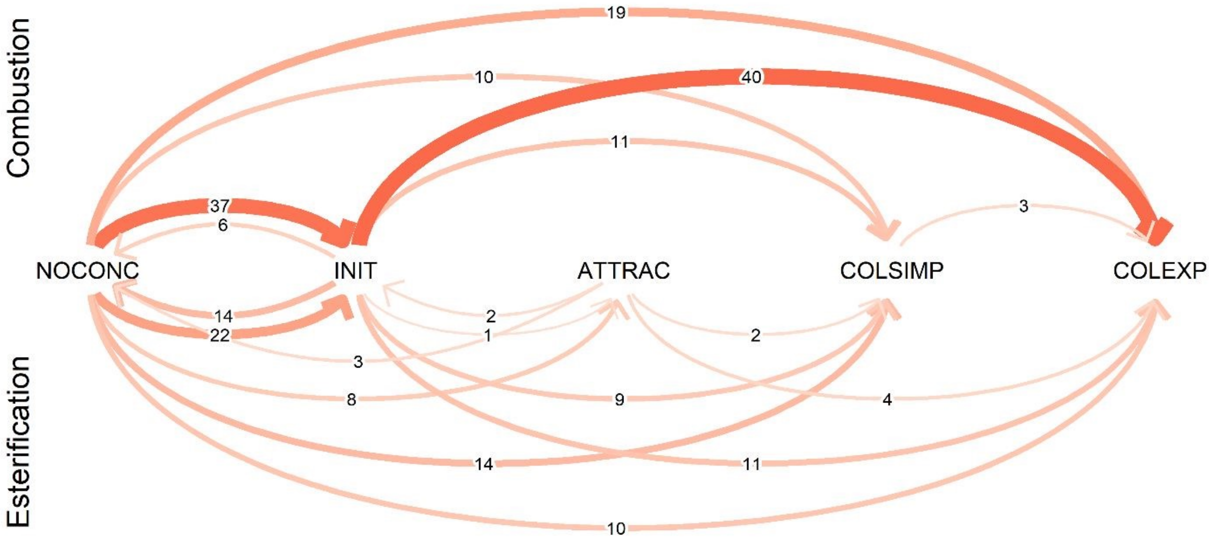

Figure 5: Transitions between sub-modes from pre to post for the reaction start (upper part: combustion, lower part: esterification). Arrowheads indicate whether transitions are progressive or regressive.

Reaction Progress. A similar trend is evident for reaction progress (blue part of Figure 4). Many learners do not activate specific conceptual knowledge in the pre-condition, chemical

View Article Online
DOI: 10.1039/D6RP00153J

bonding (BONDS) is the most salient concept. Additionally, a macroscopic focus (MACRO) also seems to take a prominent role. It reflects a primarily correct description of the reaction progress while neglecting submicroscopic considerations – thereby limiting a deeper understanding of energy changes. Although positive conceptual shifts are evident from pre to post, the macroscopic focus (MACRO) remains prominent even after instructional opportunities particularly in the context of the esterification. Figure 6 shows that in combustion, understanding bond breaking and formation (BONDS) particularly favors transitions to a comprehensive understanding of collisions (COLEXP). In esterification, many learners rely on a macroscopic focus (MACRO) as an explanatory pattern, especially if they had no to little prior concepts (NOCONC). Still, in esterification more learners move from NOCONC to BONDS than from NOCONC to COLEXP, suggesting that this level must be passed cumulatively despite new learning opportunities addressing collision theory.

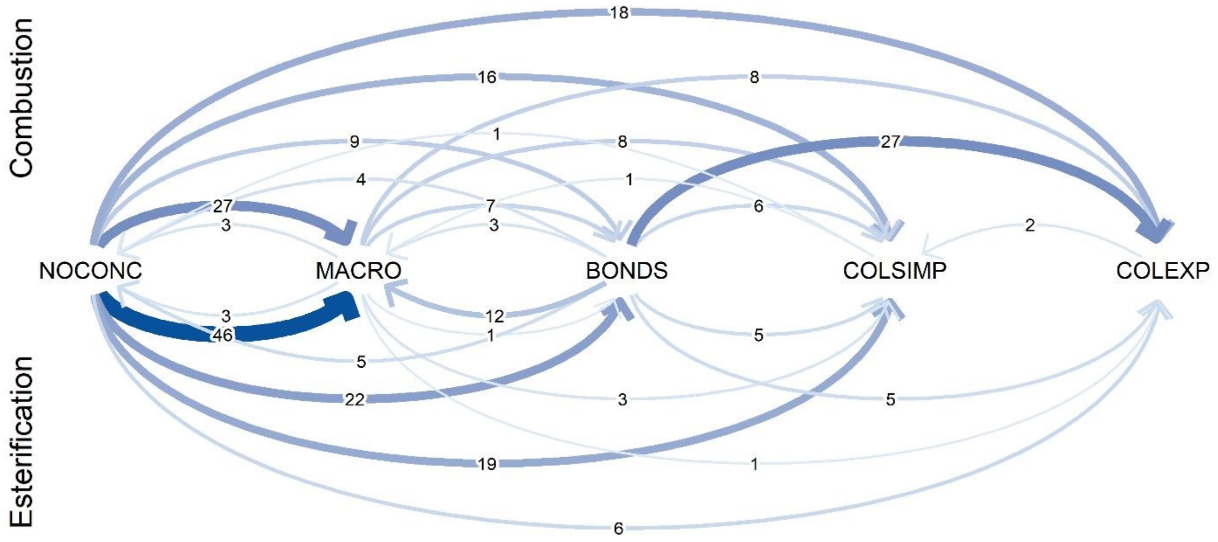

Figure 6: Transitions between sub-modes from pre to post for the reaction progress (upper part: combustion, lower part: esterification). Arrowheads indicate whether transitions are progressive or regressive.

Rection End. In the pre-condition, specific concepts are rarely activated when interpreting the end of a reaction (green part in Figure 4). This is not unexpected, as lower secondary chemistry usually treats reactions as going to completion (Van Driel, 2002). The limiting factor (LIMIT) emerges as the most salient sub-mode. In the post-condition, clear shifts appear, with

View Article Online
DOI: 10.1039/D6RP00153J

combustion remaining correctly dominated by the limiting factor, while esterification shows a more balanced distribution among no concepts (NOCONC), the limiting factor (LIMIT) and a dynamic understanding of equilibrium (DYNAMIC). For esterification, a static understanding (STATIC) is largely irrelevant as both a starting and endpoint. In contrast, LIMIT and DYNAMIC show different patterns: in combustion, transitions to LIMIT dominate, whereas transitions from NOCONC are roughly balanced between LIMIT and DYMANIC in esterification.

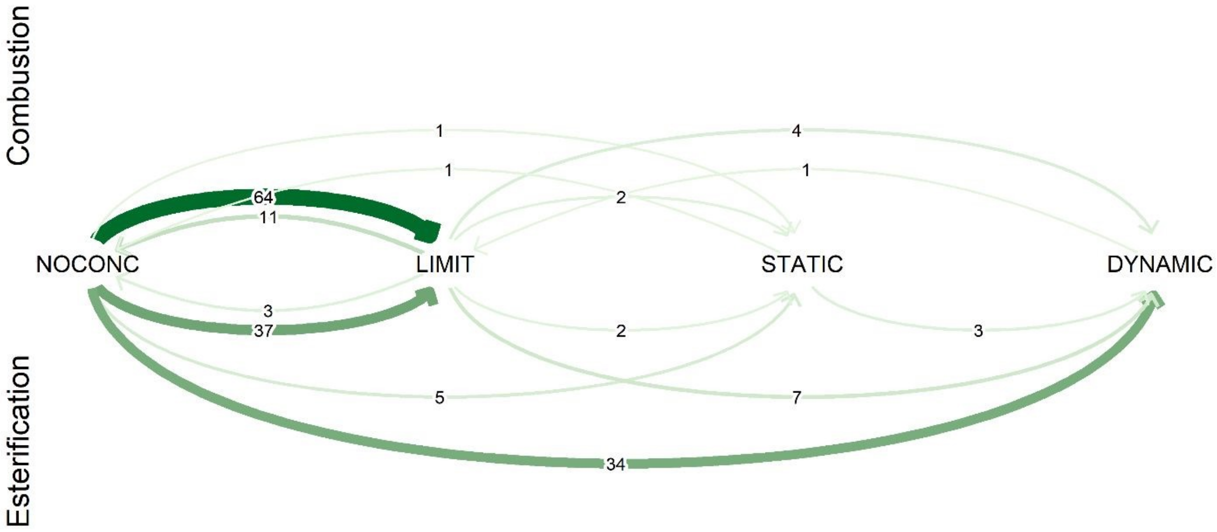

Figure 7: Transitions between sub-modes from pre to post for the reaction end (upper part: combustion, lower part: esterification). Arrowheads indicate whether transitions are progressive or regressive.

Significance of Changes in the Marginal Distributions. Overall, the tests of marginal homogeneity suggest that the marginal distributions changed substantially across both reaction examples and the total of their specific conceptual modes, thus indicating conceptual learning through instruction. However, these descriptive insights and the test of marginal homogeneity do not allow for cell-level inferences about how learners' pre-instructional sub-modes shape subsequent development (see Part III). Detailed test results for all comparisons can be found in the appendix (Table A7).

Part II - Differential Changes in Students' Conceptual Modes

We examined differences between the instructional groups (KOG and KEG) using an additive DiD approach, resulting in ATT estimators. The ATT estimator quantifies the average causal

View Article Online
DOI: 10.1039/D6RP00153J

effect of receiving instruction on both kinetics and equilibrium (KEG) compared to kinetics only (KOG). For conceptual sub-mode occupancy (0-100%), a positive ATT indicates a higher probability of occupying a sub-mode in KEG, while a negative ATT indicated the oppositive. In our report, the ATT consistently refers to the KEG group. For example, an ATT of 0.094 for the start of combustion reactions indicates that students in the KEG group had a 9.4 percentage point higher probability of being in the comprehensive understanding of collisions (COLEXP) sub-mode compared to students in the KOG group, after instruction. For clarity, we report the results of this analysis in a written form, including the respective estimates. However, we provide a full tabular summary of all ATT estimates in Table A8 in our supporting information.

Reaction Start. In accordance with our hypotheses related to RQ2, our findings indicate that, for the start of reactions, the KEG group exhibits a significantly higher occupancy of the sub-mode COLEXP across both reaction types, suggesting that these students exhibit a deeper understanding of collision theory (ATT = 0.137 with *p = .03* for the combustion reaction and 0.094 with *p = .04* for the esterification reaction, respectively).

Reaction Progress. We do not observe a significantly higher occupancy of the sub-mode COLEXP for any of the two reaction types when considering the reaction progress (ATT = -0.001 with *p = .904* and 0.040 with *p = .15*, respectively). Furthermore, our findings indicate that in KEG, particularly for the esterification reaction, a macroscopic interpretation (MACRO) is significantly more prevalent (ATT = 0.256 with *p = .001*), whereas the interpretation focusing on bond breaking and formation (BONDS) is significantly less frequent (ATT = -0.203 with *p = .007*). This development is considered unfavorable, as MACRO opposes a submicroscopic understanding of reaction processes. Ordinal regression (part III) will shed further light on conditions for development.

Reaction End. For complete combustion, no significant differences between KEG and KOG were observed regarding the end of reactions (see Table A8). Consistent with this, both groups showed similar transition patterns, with the limiting factor (LIMIT) displaying the largest gains or remaining stable over time (see Figures 4 and 7). In esterification, however, KEG students

View Article Online
DOI: 10.1039/D6RP00153J

showed a substantially higher likelihood of being classified in a dynamic understanding of equilibrium (DYNAMIC, ATT = 0.524, $p < .001$), whereas all other sub-modes were less represented. This pattern indicates differences between the groups primarily for reaction systems involving equilibrium.

Part III – Effects on Students' Sub-Modes in the Post-Condition

To further analyze conditions of conceptual development, we calculated stepwise prediction models of students' sub-modes in the post-condition. We successively incorporated general prior knowledge (last chemistry grade), specific prior knowledge (sub-modes in the pre-condition) and subject-related interest as predictor variables. In presenting the results, Figure 8 summarizes the findings of all individual regressions. These findings are subsequently explained in more detail below. The full report of all regression models and parameter estimates are available in the appendix (Section F).

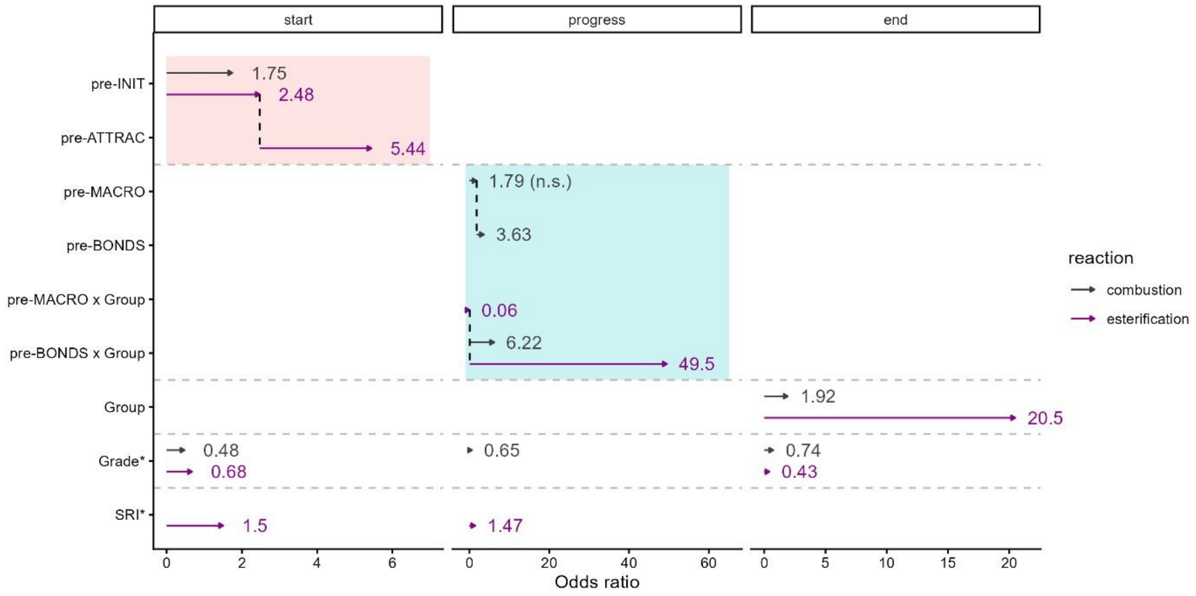

Figure 8: Summary of effects on conceptual development in terms of Odds Ratios (ORs). Only significant effects are reported (cf. appendix). ORs above 1 indicate higher probabilities of reaching sub-modes with greater explanatory power after the teaching unit, whereas ORs below 1 indicate lower probabilities. Vertical dotted lines indicate relationships between consecutive sub-modes; in these cases, the OR values indicate the cumulative effect across the sequence of sub-modes in relation to NOCONC.

The comprehensive analysis within the framework of RQ3 reveals generally consisting findings across the established analytical unit, with specific variations depending on the combination of main mode and reaction example. In contrast to our usual structure along the three main

View Article Online
DOI: 10.1039/D6RP00153J

modes (start, progress, end), we will present the findings visualized in Figure 8 according to the stepwise introduction of predictors.

Specific Prior Knowledge. A detailed examination of the start of reaction of both reaction examples reveals that students possessing at least an initial understanding (INIT) in the pre-condition have significantly improved chances of reaching a higher conceptual sub-mode after the teaching unit. For instance, already possessing INIT increases the odds for the esterification reaction by a factor of 2.5 compared to no prior concepts (NOCONC) as reference category). This finding supports the positive trend observed for INIT in the context of RQ1, as identified in the descriptive analysis of transitions between sub-modes. Furthermore, the next higher sub-mode of mutual attractions (ATTRAC) exerts an even stronger influence in case of the esterification reaction. There is a significant advantage of ATTRAC over INIT in the pre-condition, resulting in a cumulative odds ratio of 5.44 for this reaction type.

At first glance, the direct effect of sub-modes in the pre-condition appears to be lower for the progress of the reaction. While there is no significant increase in the odds ratio when comparing the macroscopic understanding (MACRO) to the reference category of no prior concepts (NOCONC), students who already possess an initial understanding of bond breaking and formation (BONDS) have a significantly improved chance of reaching a higher conceptual sub-mode after the teaching unit, with a cumulative odds ratio of 3.63 compared to NOCONC. Interestingly, the best model fit includes interactions between sub-modes in the pre-condition and group membership. The analysis of the combustion reaction indicates a significant interaction effect of BONDS and group membership, suggesting that students in the KEG group who initially reasoned at the bond level had a substantially higher chance in reaching a higher conceptual sub-mode (odds ratio of 6.22). A closer examination of the esterification reaction revealed more pronounced group-dependent differences related to BONDS. The regression model showed a negative effect of pre-MACRO – the macroscopic interpretation of esterification reactions – for KEG students.

View Article Online
DOI: 10.1039/D6RP00153J

Notably, specific prior knowledge exhibits little to no importance in relation to the end of reactions. This challenges our expectations that specific prior knowledge would be an overall key predictor. Instead, group membership emerges as the most predictive factor for the end of reactions. While this effect was already evident in the differential DiD analysis for the esterification reaction through the calculated ATTs (which is reflected here by in odds ratio of 20.5 for the KEG group compared to the KOG group), a significant group advantage also manifests in the combustion reaction (but substantially lower, with an OR of 1.92). This advantage is not directly apparent from the ATTs but rather accumulates through the significant overall category shifts.

General Prior Knowledge. Across almost all combinations of conceptual mode and reaction example, the last chemistry grade, representing general prior knowledge, emerges as the most stable predictor. The ORs for this predictor range from 0.43 to 0.74, indicating that lower (i.e., better) grades are associated with an increased likelihood of reaching a higher sub-mode category in the post-condition.

Subject-Related Interest. Subject-related interest (SRI), included as a covariate, showed significant effects only for esterification reactions. Specifically, higher SRI was associated with increased odds of reaching higher conceptual sub-modes for the start (OR = 1.50) and progress (OR = 1.47) of reactions. No significant effects were observed for reaction end of esterification reactions or for any of the three modes in the case of combustion reactions.

Discussion

This study investigated the extent to which instruction on the topics chemical kinetics and chemical equilibrium can enhance learners' conceptual understanding of chemical reactions. We analyzed these conceptual transitions from multiple perspectives, considering (I) overall changes in the distribution of leaners conceptual sub-modes before and after instruction; (II) the differential impact of instruction on one versus both of these topic areas, identifying distinct learning gains and potential conceptual difficulties; and (III) the combined consideration of the isolated effects theoretically relevant predictors.

View Article Online
DOI: 10.1039/D6RP00153J

Students' Transitions on a Conceptual Landscape: A Question of Specific Prior Knowledge?

The results of analysis steps (I) and (III) indicate that both specific prior knowledge and general prior knowledge facilitated the cumulative development of conceptual understanding of chemical reactions. While learning opportunities related to chemical kinetics and chemical equilibrium led to significant shifts in learners' conceptual modes, our complementary analyses showed that students benefit from these learning opportunities in different ways. Specifically, students with higher prior knowledge – whether reflected in pre-condition sub-modes or chemistry grades – had in most cases a significantly higher likelihood of achieving conceptual growth. This finding supports the knowledge-is-power hypothesis (Etzel et al., 2025; Hambrick and Engle, 2002), which posits that prior knowledge supports the effective use of learning opportunities. Contrary to our expectations, the relevance of specific prior knowledge did not differ as strongly from general prior knowledge. Additionally, including subject-related interest as a covariate strengthened the interpretation of prior knowledge effects by accounting for interest-related differences between students.

At the same time, the central contribution of our findings lies less in confirming the importance of prior knowledge – which is already well known from educational research – and more in demonstrating the value of assessing specific prior knowledge through conceptual sub-modes. Consistent with prior research, our analyses show that specific prior knowledge strongly predicts learning outcomes because conceptual modes build on one another (Brod, 2021; Edelsbrunner et al., 2024). More importantly, the contrast coding partially revealed the cumulative structure of explanatory levels, thereby making visible how conceptual understanding develops across interconnected modes. For example, as an intended outcome of lower secondary education, bond breaking and reformation (BONDS) generally provided a productive foundation for further conceptual development. However, our findings also highlight group-related differences in the utility of this specific prior knowledge facet, as BONDS as sub-mode in the pre-condition did not equally support all learners in advancing to higher conceptual modes. This was particularly evident in the KOG group, where BONDS supported transitions in esterification tasks less strongly. More specifically, students in KOG were unable to transfer

View Article Online
DOI: 10.1039/D6RP00153J

their knowledge to ester-related tasks as their learning environment (see Table 1, KOG students only received instruction on lesson set 1 and 2) did not explicitly provide opportunities to make such connections (Kehne, 2019).

Revealing this cumulative structure is of particular importance because it offers considerable diagnostic potential for identifying where students are progressing within the conceptual landscape and what conceptual support they may require. These implications for targeted conceptual support are discussed further in the implications section. At the same time, the sequential elaboration from lower to upper anchors of understanding highlighted in our findings demonstrates the potential of the present study to inform the design and validation of a future learning progression for upper secondary education. Overall, students who still exhibited a conceptual understanding at lower levels of reasoning by the end of lower secondary education were the least likely to benefit from new learning opportunities. This further underlines the importance of monitoring not only the quantity of students' understanding, but also their conceptual positioning and developmental trajectories.

Students' Transitions on a Conceptual Landscape: A Question of Instruction?

In line with our expectations regarding effective time-on-task (Hattie, 2010), we found that learners in the KEG group demonstrated a deeper conceptual understanding regarding the **reaction start** in both reaction examples. However, these positive effects were not fully replicated for conceptualizing the **reaction progress**, where the macroscopic understanding and interpretation of reactions (MACRO) as pre-concept proved to be an obstacle for conceptual growth, particularly within the KEG group. Specifically, students with little prior conceptual knowledge (NOCONC) were more likely to transition into MACRO as outcome, and those who started at this level were more likely to remain. This suggests a somewhat stabilizing effect of macroscopic reasoning. A more differentiated perspective reveals that KEG learners who entered the learning environment and already possessed a more sophisticated concept of bond breaking and formation (BONDS) were somewhat protected from regressing at a macroscopic level of reasoning. A critical examination of the instructional materials and contexts suggests that the contextualization of the equilibrium concept within an aroma

View Article Online
DOI: 10.1039/D6RP00153J

compound synthesis setting (see Table 1) – which frequently emphasized reaction yield, concentration trends, and concentration calculations – may have reinforced a macroscopic focus, particularly for learners who had not yet developed a submicroscopic understanding of the esterification reaction prior to the unit. With regard to developing a learning progression, this result supports the intertwined relation of instruction and learning progress as specific steps require specific instruction (Corcoran et al, 2009), questions a strictly linear sequence (Sikorski & Hammer, 2017) and underlines that not every transition between possible steps is equal, but that specific intermediate conceptions (in this case MACRO) might serve as gatekeeper. While the unit included activities that explicitly targeted a submicroscopic view of the end of reactions and their dynamic nature (Lossjew & Bernholt, 2025), aroma synthesis as a context seems to have reinforced macroscopic reasoning. At this point, stronger decontextualization and scaffolding with regard to the activation of appropriate resources may be needed (Parchmann et al., 2006, Pölloth et al., 2023). As expected, we observed the most pronounced effect for conceptualizing the **reaction end**, where KEG learners showed superior competence and an improved ability to differentiate reaction systems as they did not incorrectly overgeneralize their knowledge of the dynamic equilibrium state. This differentiation underscores that a conceptual shift between complete and incomplete reactions (Van Driel et al., 1998) generally requires explicit instruction and should be addressed even more explicitly in a revised version of the instructional material.

Limitations

On a content level, examples were limited to two typical reaction types from school chemistry education. Previous research has shown that reasoning about chemical reactions is influenced by the specific reaction example (Yan and Talanquer, 2015). While these two reactions were carefully selected to cover different mechanistic and conceptual challenges relevant to school-level chemistry, our findings cannot be fully generalized to broader competence development within the core concept of chemical reactions, particularly for other reaction types such as redox reactions or polymerization processes.

View Article Online
DOI: 10.1039/D6RP00153J

From a methodological perspective, assessing learners' conceptual modes through open-ended responses offers advantages over traditional multiple-choice assessments (Emden et al., 2018; Walpuski and Celik, 2024), as it provides deeper insights into students' reasoning processes. However, the classification of modes may also be influenced by response length and linguistic expression. In this regard, supplementary cognitive interviews could help capture students' ideas and reasoning, particularly for those who may struggle to articulate their understanding.

Although many international curricula share substantial commonalities regarding the core concept of chemical reactions, the study was conducted primarily in schools located in Northern Germany. This regional focus represents a limitation with respect to the generalizability of the findings. Similar research approaches could also be applied in international contexts to investigate students' learning about chemical reactions.

Conclusions and Implications

This study provides evidence that learning opportunities in a research-informed TLS on chemical kinetics and chemical equilibrium generally lead to positive shifts in learners' conceptual modes, supporting the idea that targeted instruction on these topics fosters conceptual progression when reasoning about chemical reactions. In addition to the effects of instruction, we could confirm the well-known role of prior knowledge in this specific learning process. Beyond its generally positive influence, the findings also provide clear indications of the cumulative structure of the core concept of chemical reaction beyond lower secondary education and suggest that learning may additionally be shaped by affective constructs such as interest. The development and validation of a more detailed learning progression, substantially informed by the present study but requiring additional, more detailed task formats, represent logical next steps toward achieving this goal. We discuss these and further implications in the following section.

View Article Online
DOI: 10.1039/D6RP00153J

Implications for Teaching and Learning

While the importance of prior knowledge as a key determinant of students' conceptual development is not a surprising finding in itself, this study particularly provided insights into the cumulative structure of the investigated conceptual sub-modes and how these can be developed through the targeted activation of prior knowledge as well as instruction focusing on broader conceptual components (chemical kinetics and equilibrium). Uncovering these structural relationships opens diagnostic potential by enabling a more differentiated identification of students' conceptual starting points and learning pathways. We therefore argue for leveraging the diagnostic potential of digitally supported learning environments to assess students' conceptual understanding prior to entering new instructional units and to address identified conceptual gaps through targeted activities. These adaptations based on pre-diagnosic activities are also referred to as macro-adaptations. In our appendix, we further elaborate on these considerations by providing concrete design implications (Figure A2). In small learning groups, teachers can already use the assessment instrument at the transition from lower to upper secondary education to evaluate specific goals of conceptual understanding in depth and to inform subsequent instructional planning. To our knowledge, at least in Germany, no systematic monitoring currently takes place at such educational transitions. Instead, responsibility largely lies with individual teachers. The proposed approach can support teachers in this diagnostic task. However, open-ended questions become difficult to implement with larger numbers of students. We briefly discuss this difficulty as part of the future research section. Beyond the macro-adaptations discussed here, digitally supported learning environments may also enable micro-adaptations, that is, task-level adaptations based on students' ongoing learning processes. This perspective was not the focus of the present study and would require further analyses at the task level to better support learners in conceptually critical tasks (e.g., Schumacher and Stern, 2023). Likewise, identifying which tasks are particularly important for conceptual growth would require additional analyses (e.g., Wyrwich et al., 2025).

View Article Online
DOI: 10.1039/D6RP00153J

Implications for Research in Conceptual Development

The assessment and analysis strategy applied in this study provides a valuable methodological framework for investigating students' conceptual development within the core concept of CR that could be extended to other questions of conceptual growth. It allows for a structured evaluation of whether learners have reached the learning objectives of lower secondary chemistry education, while also identifying conceptual transitions and reasons for divergent pathways. A key strength of this approach is its three-step design, which enables an exploration of conceptual growth from multiple perspectives. Our approach, thus, represents a first step towards the development of a learning progression by substantiating hypothetical pathways (i.e. theoretical learning progression) with empirical data and thereby preparing the ground for further LP-related analyses (e.g. Emden et al., 2018; Hadenfeldt et al., 2016).

The present study primarily focused on the mechanistic level conceptualized by Yan and Talanquer (2015), as the learning opportunities did not explicitly address causality-related reasoning regarding why chemical reactions occur. More complex structural-energetic reasoning (e.g., Asmussen et al., 2023; Braun et al., 2024) was likewise beyond the scope of the target educational level. Future research could therefore investigate how learners integrate energetic considerations into mechanistic reasoning to further elaborate a learning progression extending towards university chemistry learning.

With regard to the structure of such an LP, we formulated "developmental statements" for upper secondary education based on our findings. These statements are provided in appendix (Section I). The six statements outline how students are expected to progress in explaining matter, particle interactions, and the resulting course of chemical reactions, including how these processes are shaped by features of the reaction system. They are intended to inform further research towards the design and empirical investigation of an LP on chemical reactions in upper secondary education.

From a diagnostic perspective, evaluating complex student responses in large learning groups represents a considerable challenge for teachers. Based on the findings of our study, AI-based approaches for analyzing students' free-text responses therefore appear to be a logical next step, as they may offer promising opportunities to identify diagnostically relevant knowledge elements and support teachers' diagnostic work (e.g., Gombert et al., 2023; Bernholt et al., 2025).

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the Supplementary Information.

Acknowledgements

This work was supported by the Leibniz Foundation under the grant K365/2020. We thank Carolin Flerlage for the careful reading of the manuscript, and Clemens Frank and Jonas Bünning for their support with the coding tasks.

Notes and references

Ahtee, M., Varjola, I., 1998. Students' understanding of chemical reaction. International Journal of Science Education 20, 305–316. https://doi.org/10.1080/0950069980200304

Akkus¸, H., Kadayifc¸i, H., Atasoy, B., Geban, O., 2003. Effectiveness of instruction based on the constructivist approach on understanding chemical equilibrium concepts. Research in Science & Technological Education 21, 209–227. https://doi.org/10.1080/0263514032000127248

Anastas, P.T., Warner, J.C., 1998. Green chemistry: theory and practice. Oxford University Press, Oxford [England] ; New York.

Asmussen, G., Rodemer, M., Bernholt, S., 2023. Blooming student difficulties in dealing with organic reaction mechanisms – an attempt at systemization. Chem. Educ. Res. Pract. 24, 1035–1054. https://doi.org/10.1039/D2RP00204C

Bain, K., Towns, M.H., 2016. A review of research on the teaching and learning of chemical kinetics. Chem. Educ. Res. Pract. 17, 246–262. https://doi.org/10.1039/C5RP00176E

Banerjee, A.C., 1991. Misconceptions of students and teachers in chemical equilibrium. International Journal of Science Education 13, 487–494. https://doi.org/10.1080/0950069910130411

Bernholt, S., Lossjew, J., & Gombert, S. (2025). *From Students' Learning Trajectories to Individualized Feedback and Support*. https://doi.org/10.13140/RG.2.2.11315.69929

Bernholt, S., & Sevian, H. (2018). Learning progressions and teaching sequences – old wine in new skins? *Chemistry Education Research and Practice*, *19*(4), 989–997. https://doi.org/10.1039/C8RP90009D

Betthäuser, B.A., Bach-Mortensen, A.M., Engzell, P., 2023. A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. Nat Hum Behav 7, 375–385. https://doi.org/10.1038/s41562-022-01506-4

Boo, H., Watson, J.R., 2001. Progression in high school students' (aged 16–18) conceptualizations about chemical reactions in solution. Science Education 85, 568–585. https://doi.org/10.1002/sce.1024

Boujaoude, S.B., 1991. A study of the nature of students' understandings about the concept of burning. J Res Sci Teach 28, 689–704. https://doi.org/10.1002/tea.3660280806

Braun, I., Buchczyk, R., Graulich, N., 2024. Making the connection: The relation of structural and energetic representations in Organic Chemistry. Chemkon 31, 6–14. https://doi.org/10.1002/ckon.202300002

Brod, G., 2021. Toward an understanding of when prior knowledge helps or hinders learning. npj Sci. Learn. 6, 24. https://doi.org/10.1038/s41539-021-00103-w

Cakmakci, G., 2010. Identifying Alternative Conceptions of Chemical Kinetics among Secondary School and Undergraduate Students in Turkey. J. Chem. Educ. 87, 449–455. https://doi.org/10.1021/ed8001336

View Article Online
DOI: 10.1039/D6RP00153J

Cakmakci, G., Aydogdu, C., 2011. Designing and evaluating an evidence-informed instruction in chemical kinetics. Chem. Educ. Res. Pract. 12, 15–28. https://doi.org/10.1039/C1RP90004H

Cakmakci, G., Leach, J., Donnelly, J., 2006. Students' Ideas about Reaction Rate and its Relationship with Concentration or Pressure. International Journal of Science Education 28, 1795–1815. https://doi.org/10.1080/09500690600823490

Celik, K.N., 2022. Entwicklung von chemischem Fachwissen in der Sekundarstufe I. Validierung einer Learning Progression für die Basiskonzepte "Struktur der Materie, Chemische Reaktion und Energie" im Kompetenzbereich Umgang mit Fachwissen, Studien zum Physik- und Chemielernen. Logos Verlag, Berlin.

Childs, P.E., Hayes, S.M., O'dwyer, A., 2015. Chemistry and Everyday Life: Relating Secondary School Chemistry to the Current and Future Lives of Students, in: Eilks, I., Hofstein, A. (Eds.), Relevant Chemistry Education. SensePublishers, Rotterdam, pp. 33–54. https://doi.org/10.1007/978-94-6300-175-5_3

Corcoran, T., Mosher, F., Rogat, D., 2009. Learning Progressions in Science: An Evidence-based Approach to Reform. Consortium for Policy Research in Education.

Daniels, Z., 2008. Entwicklung schulischer Interessen im Jugendalter, Pädagogische Psychologie und Entwicklungspsychologie. Waxmann, Münster.

Duncan, R.G., Hmelo-Silver, C.E., 2009. Learning progressions: Aligning curriculum, instruction, and assessment. J Res Sci Teach 46, 606–609. https://doi.org/10.1002/tea.20316

Edelsbrunner, P.A., Schumacher, R., Hänger-Surer, B., Schalk, L., Stern, E., 2024. Preparation for future conceptual learning: Content-specific long-term effects of early physics instruction. Journal of Educational Psychology 116, 1479–1499. https://doi.org/10.1037/edu0000887

Emden, M., Weber, K., Sumfleth, E., 2018. Evaluating a learning progression on 'Transformation of Matter' on the lower secondary level. Chem. Educ. Res. Pract. 19, 1096–1116. https://doi.org/10.1039/C8RP00137E

Etzel, J.M., Heinze, A., Neumann, K., Bernholt, S., Retelsdorf, J., Köller, O., Nagy, G., 2025. What we learn in school, we learn for life: Learning opportunities as moderators of the relationship between prior knowledge and learning in post-school contexts. Contemporary Educational Psychology 80, 102336. https://doi.org/10.1016/j.cedpsych.2025.102336

García Franco, A., Taber, K.S., 2009. Secondary Students' Thinking about Familiar Phenomena: Learners' explanations from a curriculum context where 'particles' is a key idea for organising teaching and learning. International Journal of Science Education 31, 1917–1952. https://doi.org/10.1080/09500690802307730

Garnett, Patrick J., Garnett, Pamela J., Hackling, M.W., 1995. Students' Alternative Conceptions in Chemistry: A Review of Research and Implications for Teaching and Learning. Studies in Science Education 25, 69–96. https://doi.org/10.1080/03057269508560050

Goldwater, M.B., Don, H.J., Krusche, M.J.F., Livesey, E.J., 2018. Relational discovery in category learning. Journal of Experimental Psychology: General 147, 1–35. https://doi.org/10.1037/xge0000387

Gombert, S., Di Mitri, D., Karademir, O., Kubsch, M., Kolbe, H., Tautz, S., Grimm, A., Bohm, I., Neumann, K., Drachsler, H., 2023. Coding energy knowledge in constructed responses with explainable NLP models. Computer Assisted Learning 39, 767–786. https://doi.org/10.1111/jcal.12767

Graves, J.A., Fry, C., McWilliams, J.M., Hatfield, L.A., 2022. Difference-in-differences for categorical outcomes. Health Services Research 57, 681–692. https://doi.org/10.1111/1475-6773.13948

Hackling, M.W., Garnett, P.J., 1985. Misconceptions of chemical equilibrium. European Journal of Science Education 7, 205–214. https://doi.org/10.1080/0140528850070211

Hadenfeldt, J.C., Liu, X., Neumann, K., 2014. Framing students' progression in understanding matter: a review of previous research. Studies in Science Education 50, 181–208. https://doi.org/10.1080/03057267.2014.945829

Hadenfeldt, J.C., Neumann, K., Bernholt, S., Liu, X., Parchmann, I., 2016. Students' progression in understanding the matter concept: STUDENTS' PROGRESSION IN UNDERSTANDING MATTER. J Res Sci Teach 53, 683–708. https://doi.org/10.1002/tea.21312

Hambrick, D.Z., Engle, R.W., 2002. Effects of Domain Knowledge, Working Memory Capacity, and Age on Cognitive Performance: An Investigation of the Knowledge-Is-Power Hypothesis. Cognitive Psychology 44, 339–387. https://doi.org/10.1006/cogp.2001.0769

Harrison, A.G., Treagust, D.F., 2003. The Particulate Nature of Matter: Challenges in Understanding the Submicroscopic World, in: Gilbert, J.K., Jong, O., Justi, R., Treagust, D.F., Driel, J.H. (Eds.), Chemical Education: Towards Research-Based Practice, Science & Technology Education Library. Kluwer Academic Publishers, Dordrecht, pp. 189–212. https://doi.org/10.1007/0-306-47977-X_9

View Article Online
DOI: 10.1039/D6RP00153J

Hattie, J., 2010. Visible learning: a synthesis of over 800 meta-analyses relating to achievement, Reprinted. ed. Routledge, London.

Heeg, J., Bittorf, R.M., Schanze, S., 2020. Learners' conceptions about the chemical equilibrium – A systematic Review. Chemkon 27, 373–383. https://doi.org/10.1002/ckon.201900022

Hesse, J.J., Anderson, C.W., 1992. Students' conceptions of chemical change. J Res Sci Teach 29, 277–299. https://doi.org/10.1002/tea.3660290307

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, 1st ed. Wiley. https://doi.org/10.1002/0471722146

Jansen, M., Lüdtke, O., Schroeders, U., 2016. Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. Contemporary Educational Psychology 46, 116–127. https://doi.org/10.1016/j.cedpsych.2016.05.004

Johnson, P., 2002. Children's understanding of substances, Part 2: Explaining chemical change. International Journal of Science Education 24, 1037–1054. https://doi.org/10.1080/09500690110095339

Justi, R., 2003. Teaching and Learning Chemical Kinetics, in: Gilbert, J.K., Jong, O., Justi, R., Treagust, D.F., Driel, J.H. (Eds.), Chemical Education: Towards Research-Based Practice, Science & Technology Education Library. Kluwer Academic Publishers, Dordrecht, pp. 293–315. https://doi.org/10.1007/0-306-47977-X_13

Kehne, F., 2019. Analyse des Transfers von kontextualisiert erworbenem Wissen im Fach Chemie, Studien zum Physik- und Chemielernen. Logos Verlag, Berlin.

Kousathana, M., Tsaparlis, G., 2002. STUDENTS' ERRORS IN SOLVING NUMERICAL CHEMICAL-EQUILIBRIUM PROBLEMS. Chem. Educ. Res. Pract. 3, 5–17. https://doi.org/10.1039/B0RP90030C

Krnel, D., Watson, R., Glažar, S.A., 1998. Survey of research related to the development of the concept of 'matter.' International Journal of Science Education 20, 257–289. https://doi.org/10.1080/0950069980200302

Kroonenberg, P.M., Verbeek, A., 2018. The Tale of Cochran's Rule: My Contingency Table has so Many Expected Values Smaller than 5, What Am I to Do? The American Statistician 72, 175–183. https://doi.org/10.1080/00031305.2017.1286260

Kubsch, M., Czinczel, B., Lossjew, J., Wyrwich, T., Bednorz, D., Bernholt, S., Fiedler, D., Strauß, S., Cress, U., Drachsler, H., Neumann, K., Rummel, N., 2022. Toward learning progression analytics — Developing learning environments for the automated analysis of learning using evidence centered design. Front. Educ. 7, 981910. https://doi.org/10.3389/feduc.2022.981910

Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O., Reiss, K. (Eds.), 2023. PISA 2022: Analyse der Bildungsergebnisse in Deutschland. Waxmann Verlag GmbH, Münster, Germany. https://doi.org/10.31244/9783830998488

Linn, M.C., Eylon, B.-S., Rafferty, A., Vitale, J.M., 2017. Designing Instruction to Improve Lifelong Inquiry Learning. EURASIA J MATH SCI T 11. https://doi.org/10.12973/eurasia.2015.1317a

Liu, X., 2001. Synthesizing research on student conceptions in science. International Journal of Science Education 23, 55–81. https://doi.org/10.1080/09500690119778

Lossjew, J., Bernholt, S., 2024. Pressure's On: Exploring the Course of Chemical Reactions with ARDUINO and GeoGebra in a Hands-On Science Approach. J. Chem. Educ. 101, 2912–2919. https://doi.org/10.1021/acs.jchemed.4c00178

Lossjew, J., & Bernholt, S. (2025). Digital Support Custom-Made or in Moderation – Design and Evaluation of a Digitally Supported Teaching Unit on Reaction Kinetics and Chemical Equilibrium. *CHEMKON*, 32(4), 122-131. https://doi.org/10.1002/ckon.202400061.

Marsh, H.W., Trautwein, U., Lüdtke, O., Köller, O., Baumert, J., 2005. Academic Self-Concept, Interest, Grades, and Standardized Test Scores: Reciprocal Effects Models of Causal Ordering. Child Development 76, 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Mayer, R.E., 2021. Multimedia learning, Third edition. ed. Cambridge University Press, Cambridge, New York NY, Port Melbourne. https://doi.org/10.1017/978-1-316-94135-5

Méheut, M., & Psillos, D. (2004). Teaching–learning sequences: Aims and tools for science education research. *International Journal of Science Education*, *26*(5), 515–535. https://doi.org/10.1080/09500690310001614762

Ministerium für Bildung, Wissenschaft, Jugend und Kultur (Ed.), 2010. Rahmenlehrplan Naturwissenschaften für die weiterführenden Schulen in Rheinland-Pfalz. Klassenstufen 5 und 6.

Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig-Holstein (Ed.), 2019. Fachanforderungen Chemie. Allgemein bildende Schulen. Sekundarstufe I. Sekundarstufe II. Schmidt & Klaunig, Kiel.

View Article Online
DOI: 10.1039/D6RP00153J

Ministerium für Bildung, Wissenschaft, Weiterbildung und Kultur (Ed.), 2014. Lehrpläne für die naturwissenschaftlichen Fächer für die weiterführenden Schulen in Rheinland-Pfalz. Biologie. Chemie. Physik. Klassenstufen 7 bis 9/10.

Mohan, L., Chen, J., Anderson, C.W., 2009. Developing a multi-year learning progression for carbon cycling in socio-ecological systems. J Res Sci Teach 46, 675–698. https://doi.org/10.1002/tea.20314

Monteiro, L., Guerra, A., Silva, J., 2020. ESTUDO E VALIDAÇÃO DE UMA PROGRESSÃO DE APRENDIZAGEM EM CINÉTICA QUÍMICA DE ESTUDANTES DE NÍVEL SUPERIOR. Quím. Nova. https://doi.org/10.21577/0100-4042.20170647

Özmen, H., 2008. Determination of students' alternative conceptions about chemical equilibrium: a review of research and the case of Turkey. Chem. Educ. Res. Pract. 9, 225–233. https://doi.org/10.1039/B812411F

Özmen, H., Ayas, A., 2003. STUDENTS DIFFICULTIES IN UNDERSTANDING OF THE CONSERVATION OF MATTER IN OPEN AND CLOSED-SYSTEM CHEMICAL REACTIONS. Chem. Educ. Res. Pract. 4, 279–290. https://doi.org/10.1039/B3RP90017G

Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R., Ralle, B., 2006. "Chemie im Kontext": A symbiotic implementation of a context-based teaching and learning approach. International Journal of Science Education 28, 1041–1062. https://doi.org/10.1080/09500690600702512

Pölloth, B., Diekemper, D., Schwarzer, S., 2023. What resources do high school students activate to link energetic and structural changes in chemical reactions? – A qualitative study. Chem. Educ. Res. Pract. 24, 1153–1173. https://doi.org/10.1039/D3RP00068K

Scharf, M.J., List, B., 2022. A leap forward in the quest for general catalysts. Nature 610, 632–633. https://doi.org/10.1038/d41586-022-03173-x

Schettel, J. (Ed.), 2022. Konzepte zur praktischen Umsetzung der Energiewende: studentische Arbeiten aus dem Masterstudiengang Energiewirtschaftsingenieurwesen, Diskussionsbeiträge des Fachbereichs Wirtschaftsingenieurwesen der Hochschule Niederrhein / Department of Industrial Engineering, University of Applied Science Niederrhein, Krefeld. Lit, Berlin Münster.

Schneider, B., Krajcik, J., Lavonen, J., 2020. Learning Science: The Value of Crafting Engagement in Science Environments. Yale University Press. https://doi.org/10.12987/9780300252736

Schumacher, R., Stern, E., 2023. Promoting the construction of intelligent knowledge with the help of various methods of cognitively activating instruction. Front. Educ. 7, 979430. https://doi.org/10.3389/feduc.2022.979430

Schwartz, D.L., Bransford, J., Sears, D., 2005. Efficiency and innovation in transfer, in: Mestre, J.P. (Ed.), Transfer of Learning from a Modern Multidisciplinary Perspective. Information Age Publishing, Greenwich, CO, pp. 1–51.

Seçken, N., Seyhan, H.G., 2015. An Analysis of High School Students' Academic Achievement and Anxiety over Graphical Chemistry Problems about the Rate of Reaction: The Case of Sivas Province. Procedia - Social and Behavioral Sciences 174, 347–354. https://doi.org/10.1016/j.sbspro.2015.01.671

Sevian, H., Talanquer, V., 2014. Rethinking chemistry: a learning progression on chemical thinking. Chem. Educ. Res. Pract. 15, 10–23. https://doi.org/10.1039/C3RP00111C

Sikorski, T., & Hammer, D. (2017). Looking for coherence in science curriculum. *Science Education, 101*(6), 929–943. https://doi.org/10.1002/sce.21299

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.), 2024. Weiterentwickelte Bildungsstandards in den Naturwissenschaften für das Fach Chemie (MSA).

Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2020. Bildungsstandards im Fach Chemie für die Allgemeine Hochschulreife. Carl Link Verlag, Köln.

Taber, K.S., 2013. Revisiting the chemistry triplet: drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. Chem. Educ. Res. Pract. 14, 156–168. https://doi.org/10.1039/C3RP00012E

Talanquer, V., 2021. La progresión de los aprendizajes sobre la composición, estructura y transformación química de la materia. Educació Química. EduQ 4–11. https://doi.org/10.2436/20.2003.02.200

Talanquer, V., 2006. Commonsense Chemistry: A Model for Understanding Students' Alternative Conceptions. J. Chem. Educ. 83, 811. https://doi.org/10.1021/ed083p811

Treagust, D.F., Chandrasegaran, A.L., Crowley, J., Yung, B.H.W., Cheong, I.P.-A., Othman, J., 2010. EVALUATING STUDENTS' UNDERSTANDING OF KINETIC PARTICLE THEORY CONCEPTS RELATING TO THE STATES OF MATTER, CHANGES OF STATE AND DIFFUSION: A CROSS-NATIONAL STUDY. Int J of Sci and Math Educ 8, 141–164. https://doi.org/10.1007/s10763-009-9166-y

View Article Online
DOI: 10.1039/D6RP00153J

Van Driel, J.H., 2002. STUDENTS' CORPUSCULAR CONCEPTIONS IN THE CONTEXT OF CHEMICAL EQUILIBRIUM AND CHEMICAL KINETICS. Chem. Educ. Res. Pract. 3, 201–213. https://doi.org/10.1039/B2RP90016E

Van Driel, J.H., De Vos, W., Verloop, N., Dekkers, H., 1998. Developing secondary students' conceptions of chemical reactions: the introduction of chemical equilibrium. International Journal of Science Education 20, 379–392. https://doi.org/10.1080/0950069980200401

Walpuski, M., Celik, K.N., 2024. Empirical Validation of a Learning Progression in Initial Chemistry Education. CE 15, 212–237. https://doi.org/10.4236/ce.2024.152013

Woltz, D.J., Gardner, M.K., Bell, B.G., 2000. Negative transfer errors in sequential cognitive skills: Strong-but-wrong sequence application. Journal of Experimental Psychology: Learning, Memory, and Cognition 26, 601–625. https://doi.org/10.1037/0278-7393.26.3.601

Wyrwich, T., Kubsch, M., Drachsler, H., & Neumann, K. (2025). Tracking students' progression in developing understanding of energy using AI technologies. *Physical Review Physics Education Research*, *21*(1), 010152. https://doi.org/10.1103/PhysRevPhysEducRes.21.010152

Yan, F., Talanquer, V., 2015. Students' Ideas about How and Why Chemical Reactions Happen: Mapping the conceptual landscape. International Journal of Science Education 37, 3066–3092. https://doi.org/10.1080/09500693.2015.1121414

Yao, J., Guo, Y., 2018. Validity evidence for a learning progression of scientific explanation. J Res Sci Teach 55, 299–317. https://doi.org/10.1002/tea.21420

Supporting Information for

View Article Online
DOI: 10.1039/D6RP00153J

Tracking Student Learning Across a Conceptual Landscape: Transitions and Differential Changes in Conceptual Modes During a Unit on Chemical Kinetics and Equilibrium

Jannik Lossjew[a] and Sascha Bernholt[a]

[a]Leibniz-Institute for Science and Mathematics Education, Olshausenstr. 62, D-24118 Kiel, Schleswig-Holstein, Germany

View Article Online
DOI: 10.1039/D6RP00153J

A. Competence goals of lower secondary education in both federal states (Schleswig-Holstein, Rhineland-Palatinate) with regard to the six developmental statements

Since we adhered to the exact wording of the respective curricula, in one case competence goals are formulated as operationalized learning objectives and in the second case more strongly as a description of content.

Table A1: Expected learning-related baseline based on the regional curricula with reference to the six formulated developmental statements.

Developmental Statement	Schleswig-Holstein	Rhineland-Palatinate
I	• explain the composition of substances and mixture using a particle model • describe and explain changes in states of matter using a particle model • apply their prior knowledge of the structure of matter to predict possible chemical reactions • explain specific properties of molecular substances using intermolecular interactions	• particle concepts are introduced and expanded: substances consist of particles (1), the particles of a pure substance are all identical (2), and particles are in motion (3) • polarity of compounds is determined by their molecular structure • electron pair bonds and intermolecular interactions determine the spatial structures of plastics
II	• describe activation energy as the energy required to bring substances into a reactive state	• -
III	• describe and explain chemical bonding in molecules • explain changes in chemical reactions at the atomic level • explain the energy balance of chemical reactions in simple terms with the breaking and formation of chemical bonds	• the bonds between atoms in hydrogen, oxygen, water and methane (as well as other hydrocarbons) are based on shared electron pairs • matter and energy transformations are modeled based on changes in molecules and electron pair bonds • matter transformations are modeled based on changes in particles and their bonds
IV	• explain different states of matter of a substance using the relationship between the kinetic energy of particles and temperature • describe activation energy as the energy required to bring substances into a reactive state • describe the effect of a catalyst on the activation energy	• the states of matter of water are explained using simple particle models • the change in energy carrier (energy release) [during chemical reactions] is noticeable through heating, movement or light • chemical reactions are controlled by varying the reaction conditions • catalysts are typically used to lower the activation energy

View Article Online
DOI: 10.1039/D6RP00153J

V	• -	• the formation and decomposition of a metal oxide are in principle reversible • chemical reactions in an accumulator are reversible
VI	• describe the effect of a catalyst on the activation energy	• catalysts are typically used to lower the activation energy

B. Coding Manual for assessing conceptual modes of the esterification reaction

View Article Online
DOI: 10.1039/D6RP00153J

Table A2: Overview of modes and sub-modes for the esterification reaction. (*) Questions were translated from German. (**) As initial factors, we included both the mention of single compounds as an initiator and external factors. The table continues on the following page.

Mode	Open Text Item*	Sub-modes	Description	Example	Cohens κ
Start (MS)	1) Describe at the submicroscopic level how acetic acid and ethanol molecules behave when the liquids are combined. 2) Explain why combining the two liquids directly causes the reaction to start.	Advanced Collision (**COLEXP**)	Students use the concept of random effective collisions and resulting (electronic) interactions to explain the reaction start.	The alcohol and acid molecules are always in motion. When these collide with sufficient kinetic energy, this initiates the start of the reaction. The fact that the activation energy of certain particle pairs can already by exceeded at room temperature also plays a role here.	0.71 (pre); 0.81 (post)
		Simple Collision (**COLSIMP**)	Students use the concept of random collisions without further interactions.	As soon as two molecules collide successfully, a reaction occurs.	
		Mutual Attractions (**ATTRAC**)	Students use the concept of mutual attraction to explain the start of reaction.	Polarization of the alcohol and acid molecules causes them to approach, leading to their reaction.	
		Initial Factor** (**INIT**)	Students use the concept of an initiating substance (reactivity) or name external factor as reason for the reaction start.	During an esterification, acid molecules a particularly reactive and start the reaction.	
		No Concept (**NOCONC**)	No concepts are used to explain the reaction start.	The two substances react to form an ester and water.	
Progress (MP)	3) Describe how you imagine the reaction at the particle level to proceed.	Advanced Collision (**COLEXP**)	Students use a combination of different interactions (continued collisions – including product particles, bond breaking/forming, escape/remaining of a substance in the system) to explain the progress.	Ethanol and acetic acid molecules continuously collide. If this happens (by chance) with sufficient kinetic energy, a bond is formed (between the hydroxyl group and the carboxyl carbon atom), followed by the cleavage of a stable water molecule. Product molecules become more.	0.81 (pre); 0.85 (post)
		Simple Collision (**COLSIMP**)	Students use the concept of perpetual collisions without describing further interactions.	The molecules continue to collide and thereby react, but they can also rebound after colliding. Product molecules become more.	
		Bond breaking and reformation (**BONDS**)	Students only use the concept of bond breaking/forming to explain the progress.	New bonds are formed between ethanol molecules and acetic acid molecules, followed by bond cleavage with the elimination of a water molecule.	
		Macroscopic change of substances (**MACRO**)	Students describe the course of the reaction from a substance-based perspective (e.g. increase in products/decrease in reactants)	More and more ester is formed.	
		No concept (**NOCONC**)	No concepts are used to explain the progress.	Acetic acid is more acidic than ethanol.	

View Article Online
DOI: 10.1039/D6RP00153J

Mode	Open Text Item	Sub-modes	Description	Example	Cohens κ
End (ME)	4) Describe how you imagine the end of the reaction at the particle level. Which particles are still present at this point? 5) Explain how the reaction ends at this point – are processes still detectable at the macroscopic or submicroscopic level?	Dynamic Equilibrium (**DYNAMIC**)	Students explain that reactions do not actually end, but instead reach a state of dynamic equilibrium.	The ratio between reactants and products no longer changes (macroscopic perspective). However, both the forward and reverse reactions continue to occur constantly at equal reaction rates, so that no net compositional change is observable.	0.85 (pre); 0.91 (post)
		Static Equilibrium (**STATIC**)	Students explain the and of the reaction by the achievement of a static equilibrium.	A state of equilibrium is established. Once this state is reached, neither the forward nor the reverse reaction occurs, which is why the ratio of reactants to products remains constant.	
		Limiting Component (**LIMIT**)	Students explain the reaction end by the presence of a limiting factor and/or by a complete reaction of the reactants.	The reaction is complete when all acetic acid and ethanol molecules have fully reacted.	
		No concept (**NOCONC**)	No concepts are used to explain the reaction end.	A colorless liquid is still visible in the flask.	

View Article Online
DOI: 10.1039/D6RP00153J

C1. Rationale for the Choice of the test for marginal homogeneity

Literature suggests, that the “classic” Chi^2 test for independence should not be used in paired samples (Agresti, 2002). We therefore discussed different test that would fit our data. In general, the McNemar-Bowker test replaces the "classic" Chi^2 test for contingency tables that exceed a 2×2 design. However, this test becomes unreliable when many cells have a frequency of zero, as (I) it cannot be computed globally in such cases, and (II) cell-level statements about symmetry should not be interpreted when cells in the pre-condition are unpopulated. For this reason, we applied a test of marginal homogeneity, which allows conclusions about overall changes in the marginal distributions of the transition matrix. This test, however, limits cellwise interpretations and – in our case – does not allow for conclusions about dependencies between pre- and post-sub-modes. To examine such dependencies, we used ordinal regression analysis.

C2. Theoretical Rationale for the Choice of the Additive Approach and Details on ATT Estimation

Applying the classical DiD framework to categorical outcomes such as conceptual modes presents challenges, as traditional approaches typically assume continuous outcomes relying on linear relationships (Graves et al., 2022). Graves et al. (2022) extend the classical DiD framework to research settings with categorical outcomes, explicitly addressing how treatment effects, i.e. taught contents, can be identified under different assumptions about transitions across outcome states and marginal outcome distributions.

A key distinction in their approach is the comparison between additive and multiplicative assumptions when modeling changes over time. The additive approach assumes that, in the absence of treatment, pre-existing differences between groups would remain constant over time – referred to as fixed group differences. In contrast, the multiplicative approach assumes that treatment and comparison group share a common transition structure. This means that while treatment may accelerate learning, both groups will eventually converge to the same long-term distribution, regardless of whether they receive treatment or not. Both scenarios are represented in figure 1, where the intervention effect of both approaches is shown as difference

View Article Online
DOI: 10.1039/D6RP00153J

($d_2 - d_3$) between group differences d_2 (expected group difference in the post-condition without intervention) and d_3 (expected group difference in the post-condition with intervention). Under the additive assumption, this intervention effect is defined as deviation from the fixed differences which the two groups would have maintained over time without any sort of intervention, i.e. any additional change in the treated group relative to the comparison group is attributed to the intervention. Under the multiplicative assumption, both groups follow the same transition patterns, progressing toward a shared stationary distribution. Without intervention, the gap between the groups would change over time (no fixed effects). The intervention, however, disrupts this natural convergence by accelerating the treated group's progress, thereby widening the gap. This distinction is crucial because it determines how treatment effects are identified and whether differences in outcome distributions are attributed to the intervention or to pre-existing structural differences.

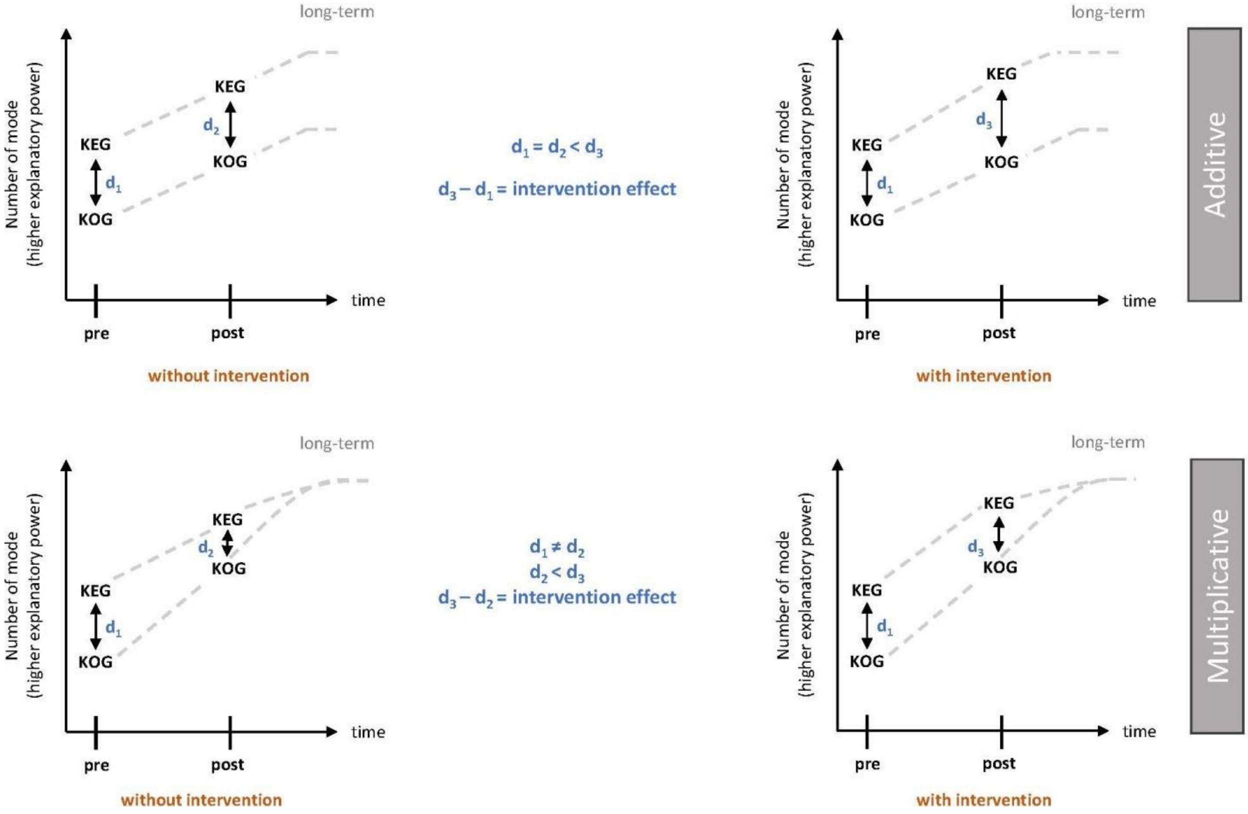

Figure A1: Comparison of theoretical developmental lines following either the additive approach (upper part of the figure) or the multiplicative approach (lower part of the figure).

In both cases, however, an ATT (Average Effect of the Treatment on the Treated) estimator is calculated, capturing the causal impact of the intervention by comparing the changes in

View Article Online
DOI: 10.1039/D6RP00153J

outcomes between the treatment and comparison groups over time. Following the framework of Graves et al. (2022), when treatment and comparison groups share the same baseline distributions, both the additive and multiplicative ATT estimators produce identical treatment effect estimates. However, when baseline distributions differ, the choice between these two options must be guided by theoretical considerations. In our study, students were naturally assigned to one of the two instructional conditions: (I) KOG received instruction only on chemical kinetics and (II) KEG received instruction on both chemical kinetics and chemical equilibrium. As in our case the two groups already exhibited differences in conceptual mode distributions at baseline, we explicitly choose the additive approach based on the following theory-driven reasoning: prior knowledge is a well-established predictor of learning outcomes in the sense of knowledge stability, with numerous studies indicating that differences in initial knowledge levels tend to remain stable over time when learning opportunities are held constant (Simonsmeier et al., 2022). Given that prior knowledge plays a key role in shaping individual progressions, we argue that pre-existing differences between groups would have remained constant in the absence of an intervention. Unlike more basic developmental processes, where individuals typically reach similar end states, conceptual learning processes in academic contexts often do not adhere to shared transition patterns or lead to stationary distributions. Thus, any deviation from this fixed difference should be attributed to the instructional intervention. Consequently, we adopt the additive approach, as it aligns with our theoretical expectations and empirical findings on the stability of knowledge differences over time. To quantify the effect of additional instruction on chemical equilibrium concretely, we measure the probability of students occupying each conceptual mode before and after instruction in both groups. This allows us to estimate how the probability of being in a given mode changes with each group over instructional time and then compare these changes across groups. Mathematically, the key idea that any change in the treatment group that exceeds the change in the control group can be attributed to intervention, is expressed as follows:

$$\pi = \left(P_{post}^{KEG} - P_{pre}^{KEG}\right) - (P_{post}^{KOG} - P_{pre}^{KOG})$$

where P_{pre} and P_{post} represent the probabilities of a student being in a specific conceptual mode before and after instruction, for groups KOG and KEG, respectively. The estimator π corresponds to the aforementioned ATT, as it measures the additional effect of receiving instruction on chemical equilibrium for students in the treatment group (KEG) compared to the control group (KOG). One can see, that under these assumptions, fixed differences without intervention would lead to an ATT that equals 0, representing the stability of knowledge differences over time. Again, following Graves et al. (2022), we assess the statistical significance on the estimated ATT using permutation tests. Therefore, we implement a permutation-based inference procedure. The treatment condition (KOG, KEG) is randomly reassigned across students while preserving the original group sizes. For each permutation, the ATT estimator is recomputed, generating an empirical null distribution under the assumption of no treatment effect. The observed ATT is then compared to this distribution, and the proportion of permuted estimates that are equal to or exceed the observed ATT serves as the p-value. Since our hypotheses specify a directional effect, in most cases we compute one-sided p-values (see results for specifications). For a comprehensive description of the full ATT-procedure, interested readers are referred to Graves et al. (2022).

C3. Details on our ordinal regression analyses

Multicollinearity. To ensure statistical validity of our ordinal regression models, we first conducted a multicollinearity check for all predictors. We applied the Generalized Variance Inflation Factor (GVIF) for categorial variables and the traditional VIF for continuous predictors (Fox and Monette, 1992). The interpretation of VIF (or squared GVIF) values followed established threshold criteria, where values between 1 and 5 indicated mild but negligible multicollinearity, values between 5 and 10 suggested considerable multicollinearity that is noteworthy but not necessarily problematic and values exceeding 10 signaled severe multicollinearities, requiring intervention. In order to maintain a model’s stability, we removed variables if their VIF/squared GVIF were substantially above 5 and suggested a strong redundancy with another predictor. In such cases, we identified the variable combination responsible for the collinearity and removed the less theoretically relevant predictor.

View Article Online
DOI: 10.1039/D6RP00153J

Model-building and selection. The model-building process started with a null model, which included only the intercept, and progressively incorporated relevant predictors in a predefined order: first, the last grade in chemistry was added to capture general prior knowledge, followed by learners' prior conceptual modes, representing specific prior knowledge. Finally, interest was included as an affective variable that is well known to be linked to academic achievement. At each step, we tested whether the inclusion of a new predictor significantly improved model fit by conducting likelihood-ratio tests, which conceptually resemble ANOVA as they assess model improvement by comparing changes in deviance. While a lower deviance generally indicates better model fit, improvements that are too small may not reach statistical significance, suggesting that an additional predictor does not substantially enhance explanatory power (Hosmer and Lemeshow, 2000). In this case the more parsimonious model was retained for the specific combination of main mode and reaction example.

Proportional Odds Assumption and Partial Proportional Odds Models. Following model selection, we assessed the proportional odds assumption for each model using the Brant test (Brant, 1990), which evaluates whether the assumption holds consistently across all levels of the ordinal dependent variable. In cases where the test indicated a potential violation, we estimated a partial proportional odds (PPO) model to examine whether the effects of the predictor in question varied substantially across cut points. If the PPO model revealed only minor deviations, with effects remaining in the same direction across thresholds, we followed Agresti (2002) in opting for the more parsimonious proportional odds model, as such small departures are often of limited practical relevance and may reflect sampling variability (in our case for example low frequencies in some outcome categories) rather than substantive differences. However, when the PPO model indicated substantial divergence in threshold-specific effects, suggesting meaningful differences how the respective predictor operates at different levels of the outcome, we retained the more flexible PPO model to adequately capture these variations (Christensen, 2019).

Statistical significance and Interpretation of Odd Ratios. Since our hypotheses were directional, we computed one-tailed p-values for all respective predictors. To enhance the

DOI: 10.1039/D6RP00153J

interpretability of our findings, we report Odds Ratios (ORs) alongside the regression coefficients. This allows for a more intuitive understanding of the magnitude of predictor effects in an ordinal regression model. For example, an OR of 2.5 for a continuous variable implies that for each unit increase in this predictor, the odds of being classified in a higher conceptual mode were 2.5 times greater, holding all other variables constant. For ordinal predictors, the OR reflected the odds of being in a higher outcome category compared to the respective reference category. ORs are particularly relevant when examining the report of all regression models provided in section E of this appendix.

View Article Online
DOI: 10.1039/D6RP00153J

D. Comparison of descriptive statistics of the predictor variable grade and subject-related interest (SRI) between the full and filtered (complete cases) dataset

Table A3: Descriptive summary of the predictors last chemistry grade and subject related interest, without restricting to complete cases.

Variable	**Total (N)**	**Mean**	**SD**	**Min**	**Max**
Grade	232	2.58	1.21	1.00	6.00
SRI	232	2.55	0.72	1.00	4.00

Of the 245 students from classes that participated in both, pre- and post-test, 232 provided data for both variables. The descriptive statistics indicate that the filtered dataset for ordinal regressions (N = 150) closely reflects the full sample, with only minimal differences in the mean value of grade (Δgrade = 0.06), suggesting no substantial bias toward selectively including only high-achieving learners.

View Article Online
DOI: 10.1039/D6RP00153J

E. Transition matrices

Reaction Start

Table A4: Pre- and post-distribution (marginals) and transitions between sub-modes from pre to post (transition-related and overall pre- and post-distribution percentages in parentheses) for the conceptual mode start of the reaction (MS).

Post / **Pre**	NOCONC	INIT	ATTRAC	COLSIMP	COLEXP	**Pre (total)**
COMBUSTION						
NOCONC	20 (23.3%)	37 (43.0%)	0 (0.0%)	10 (11.6%)	19 (22.1%)	**86 (47.0%)**
INIT	6 (6.4%)	37 (39.4%)	0 (0.0%)	11 (11.7%)	40 (42.5%)	**94 (51.4%)**
ATTRAC	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0 (0.0%)**
COLSIMP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	3 (100.0%)	**3 (1.6%)**
COLEXP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0 (0.0%)**
Post (total)	**26 (14.2%)**	**74 (40.4%)**	**0 (0.0%)**	**21 (11.5%)**	**62 (33.9%)**	**183 (100%)**
ESTERIFICATION						
NOCONC	62 (53.4%)	22 (19.0%)	8 (6.9%)	14 (12.1%)	10 (8.6%)	**116 (63.4%)**
INIT	14 (25.9%)	19 (35.1%)	1 (1.9%)	9 (16.7%)	11 (20.4%)	**54 (29.5%)**
ATTRAC	3 (23.1%)	2 (15.4%)	2 (15.4%)	2 (15.4%)	4 (30.7%)	**13 (7.1%)**
COLSIMP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0 (0.0%)**
COLEXP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0 (0.0%)**
Post (total)	**79 (43.1%)**	**43 (23.5%)**	**11 (6.0%)**	**25 (13.7%)**	**25 (13.7%)**	**183 (100%)**

View Article Online
DOI: 10.1039/D6RP00153J

Reaction Progress

Table A5: Pre- and post-distribution (marginals) and transitions between sub-modes from pre to post (transition-related and overall pre- and post-distribution percentages in parentheses) for the conceptual mode progress of the reaction (MP). All abbreviations refer to the sub-modes introduced in table 2.

Post **Pre**	NOCONC	MACRO	BONDS	COLSIMP	COLEXP	**Pre (total)**
COMBUSTION						
NOCONC	27 (27.8%)	27 (27.8%)	9 (9.3%)	16 (16.5%)	18 (18.6%)	**97** **(53.0%)**
MACRO	3 (9.4%)	6 (18.7%)	7 (21.9%)	8 (25.0%)	8 (25.0%)	**32** **(17.5%)**
BONDS	4 (8.5%)	3 (6.4%)	7 (14.9%)	6 (12.8%)	27 (57.4%)	**47** **(25.7%)**
COLSIMP	1 (33.3%)	1 (33.3%)	0 (0.0%)	1 (33.3%)	0 (0.0%)	**3** **(1.6%)**
COLEXP	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (50.0%)	2 (50.0%)	**4** **(2.2%)**
Post (total)	**35** **(19.1%)**	**37** **(20.2%)**	**23** **(12.6%)**	**33** **(18.0%)**	**55** **(30.1%)**	**183** **(100%)**
ESTERIFICATION						
NOCONC	45 (32.6%)	46 (33.3%)	22 (15.9%)	19 (13.8%)	6 (4.4%)	**138** **(75.4%)**
MACRO	3 (23.1%)	5 (38.5%)	1 (7.7%)	3 (23.1%)	1 (7.7%)	**13** **(7.1%)**
BONDS	5 (15.6%)	12 (37.5%)	5 (15.6%)	5 (15.6%)	5 (15.6%)	**32** **(17.5%)**
COLSIMP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0** **(0.0%)**
COLEXP	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**0** **(0.0%)**
Post (total)	**53** **(29.0%)**	**63** **(34.3%)**	**28** **(15.3%)**	**27** **(14.8%)**	**12** **(6.6%)**	**183** **(100%)**

View Article Online
DOI: 10.1039/D6RP00153J

Reaction End

Table A6: Pre- and post-distribution (marginals) and transitions between sub-modes from pre to post (transition-related and overall pre- and post-distribution percentages in parentheses) for the conceptual mode end of the reaction (ME). Since one person in the full sample did not answer the last question, the sample size here is reduced to N = 182. All abbreviations refer to the sub-modes introduced in table 2.

Post **Pre**	NOCONC	LIMIT	STATIC	DYNAMIC	**Pre (total)**
COMBUSTION					
NOCONC	40 (38.1%)	64 (61.0%)	1 (0.9%)	0 (0.0%)	**105 (57.8%)**
LIMIT	11 (14.7%)	58 (77.3%)	2 (2.7%)	4 (5.3%)	**75 (41.0%)**
STATIC	1 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	**1 (0.6%)**
DYNAMIC	0 (0.0%)	1 (100.0%)	0 (0.0%)	0 (0.0%)	**1 (0.6%)**
Post (total)	**52 (28.6%)**	**123 (67.6%)**	**3 (1.6%)**	**4 (2.2%)**	**182 (100%)**
ESTERIFICATION					
NOCONC	71 (48.3%)	37 (25.2%)	5 (3.4%)	34 (23.1%)	**147 (80.8%)**
LIMIT	3 (13.0%)	11 (47.8%)	2 (8.7%)	7 (30.5%)	**23 (12.6%)**
STATIC	0 (0.0%)	0 (0.0%)	1 (25.0%)	3 (75.0%)	**4 (2.2%)**
DYNAMIC	0 (0.0%)	0 (0.0%)	0 (0.0%)	8 (100.0%)	**8 (4.4%)**
Post (total)	**74 (40.6%)**	**48 (26.4%)**	**8 (4.4%)**	**52 (28.6%)**	**182 (100%)**

View Article Online
DOI: 10.1039/D6RP00153J

F. Full report of the tests for marginal homogeneity

Table A7: Overview of the results of the tests for marginal homogeneity.

	χ^2	df	p
Combustion			
Start (MS)	96.59	4	< .001
Progress (MP)	90.87	4	< .001
End (ME)	38.38	3	< .001
Esterification			
Start (MS)	52.72	4	< .001
Progress (MP)	90.42	4	< .001
End (ME)	73.13	3	< .001

View Article Online
DOI: 10.1039/D6RP00153J

G. Full report of ATT estimates

Table A8: Overview of ATT-estimators. The estimator values are interpreted for KEG (who has been provided with extended instructional options) in relation to KOG. A negative estimator indicates a lower probability of occupying the respective category, while a positive estimator suggests a higher probability. All p-values have been calculated using a permutation-bases approach

	Combustion Reaction					**Esterification Reaction**				
Start (MS)	**Sub-mode**	**ATT**	**P**	**Lower CI**	**Upper CI**	**Sub-mode**	**ATT**	**P**	**Lower CI**	**Upper CI**
	NOCONC	-0.032	.365	.359	.371	NOCONC	-0.008	.476	.472	.478
	INIT	-0,161	.068	.063	.073	INIT	-0.119	.105	.099	.111
	ATTRAC	-	-	-	-	ATTRAC	-0.030	.292	.285	.299
	COLSIMP	0.059	.135	.129	.141	COLSIMP	0.047	.192	.185	.199
	COLEXP	**0.137**	**.034**	**.030**	**.038**	**COLEXP**	**0.094**	**.038**	**.035**	**.042**
Progress (MP)	**Sub-mode**	**ATT**	**p**	**Lower CI**	**Upper CI**	**Sub-mode**	**ATT**	**p**	**Lower CI**	**Upper CI**
	NOCONC	-0.086	.195	.188	.202	NOCONC	-0.150	.061	.056	.067
	MACRO	0.123	.159	.149	.170	**MACRO**	**0.256**	**.001**	**< .001**	**.002**
	BONDS	-0,056	.265	-258	.272	**BONDS**	**-0.203**	**.007**	**.005**	**.009**
	COLCIMP	0.028	.334	.327	.340	COLCIMP	0.053	.169	.162	.175
	COLEXP	-0.001	.904	.895	.912	COLEXP	0.040	.153	.146	.159
End (ME)	**Sub-mode**	**ATT**	**p**	**Lower CI**	**Upper CI**	**Sub-mode**	**ATT**	**p**	**Lower CI**	**Upper CI**
	NOCONC	-0.137	.088	.082	.093	**NOCONC**	**-0.155**	**.054**	**.051**	**.059**
	LIMIT	0.133	.194	.183	.205	**LIMIT**	**-0.287**	**< .001**	**< .001**	**.001**
	STATIC	0.006	.822	.811	.832	**STATIC**	**-0.082**	**.018**	**.016**	**.021**
	DYNAMIC	-0.002	.867	.857	.876	**DYNAMIC**	**0.524**	**< .001**	**< .001**	**< .001**

View Article Online
DOI: 10.1039/D6RP00153J

H. Full report of all estimated models

MS Combustion

Table A9: Estimated coefficients and significance levels for the ordinal regression model for the start of the combustion reaction (MS Combustion).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
pre-INIT	**0.558 ***	**0.333**	**1.747 [1.010; ∞]**	**-**	**2.809**
Group	0.348 (n.s.)	0.340	1.417 [0.810; ∞]	-	1.046
Grade	**-0.742 *****	**0.160**	**0.476 [0; 0.619]**	**-**	**21.492**
Threshold Coefficients					
0|1	-3.760	0.623	-	-	36.350
1|3	-1.265	0.527	-	-	5.753
3|4	-0.679	0.519	-	-	1.701

As pre-ATTRAC contained no observations in the pre-condition and pre-COLSIMP had too few observations to yield reliable estimates, only pre-INIT was included as pre-sub-mode in the model. As only one sub-mode of the ordinally structured predictor (pre-INIT) is included in the model, variant (a) with NOCONC as the reference category and variant (b) with contrast coding yield identical results. Therefore, we only report variant (a). The inclusion of SRI did not enhance model fit, leading to its exclusion as an explanatory variable in the final reported model. Interaction terms were deemed unnecessary based on a pre-check of separate regressions for the two groups (KOG, KEG), which revealed no differences in the strength or direction of the predictor effects. Model fit indices were as follows: log-likelihood (LLH) = -165.48, null model log-likelihood (LLH_{Null}) = -186.42, and likelihood-ratio test G^2 = 41.87 (p < .001). Nagelkerke's pseudo-R^2 was 0.266. Significance levels: *** p ≤ .001, ** p ≤ .01, * p ≤ .05.

MS Esterification

Table A10: Estimated coefficients and significance levels for the ordinal regression model for the start of the esterification reaction (MS Esterification).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
pre-INIT	**0.906 ****	**0.351**	**2.476 [1.391; ∞]**	**-**	**6.682**
pre-ATTRAC	**1.995 *****	**0.626**	**7.351 [2.627; ∞]**	**-**	**10.167**
SRI	**0.408 ***	**0.238**	**1.504 [1.016; ∞]**	**-**	**2.939**
Group	0.408 (n.s.)	0.367	1.504 [0.821; ∞]	-	1.231
Grade	**-0.391 ***	**0.183**	**0.676 [0; 0.913]**	**-**	**4.589**
SRI x Group	-0.595 (n.s.)	0.356	-	0.552 [0.275; 1.108]	2.791

View Article Online
DOI: 10.1039/D6RP00153J

Threshold Coefficients					
0|1	-0.809	0.504	-	-	2.572
1|2	0.249	0.502	-	-	0.243
2|3	0.559	0.506	-	-	1.219
3|4	1.457	0.525	-	-	7.590
Variant (b) – Sequential contrast coding					
pre-INIT	**0.906 ****	**0.351**	**2.476 [1.391; ∞]**	**-**	**6.682**
pre-ATTRAC	**1.088 ***	**0.650**	**2.969 [1.019; ∞]**	**-**	**2.799**
SRI	**0.408 ***	**0.238**	**1.504 [1.016; ∞]**	**-**	**2.939**
Group	0.408 (n.s.)	0.367	1.504 [0.821; ∞]	-	1.231
Grade	**-0.391 ***	**0.183**	**0.676 [0; 0.913]**	**-**	**4.589**
SRI x Group	-0.594 (n.s.)	0.355	-	0.552 [0.275; 1.108]	**2.791**
Threshold Coefficients					
0|1	-1.776	0.533	-	-	11.084
1|2	-0.718	0.517	-	-	1.928
2|3	-0.408	0.517	-	-	0.624
3|4	0.490	0.526	-	-	0.864

Model fit indices were as follows: log-likelihood (LLH) = -196.63, null model log-likelihood (LLH_{Null}) = -213.07, and likelihood-ratio test $G^2 = 32.86$ ($p < .001$). Nagelkerke's pseudo-R^2 was 0.209. Except for the interaction term, we have theory-driven directional assumptions for predictors, so one-sided CIs are reported. Significance levels: *** p ≤ .001, ** p ≤ .01, * p ≤ .05.

View Article Online
DOI: 10.1039/D6RP00153J

MP Combustion

Table A11: Estimated coefficients and significance levels for the ordinal regression model for the progress of the combustion reaction (MP Combustion).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
pre-MACRO	0.580 (n.s)	0.435	1.786 [0.845; ∞]	-	1.623
pre-BONDS	**1.119 ****	**0.519**	**3.303 [1.406; ∞]**	**-**	**5.295**
Grade	**-0.428 *****	**0.152**	**0.652 [0; 0.838]**	**-**	**7.885**
pre-MARCO x Group	-0.523 (n.s.)	0.832	0.592 [0; 2.330]	-	0.396
pre-BONDS x Group	1.305 (n.s.)	0.883	3.687 [0.862; ∞]	-	2.181
Threshold Coefficients					
0|1	-1.932	0.535	-	-	13.068
1|2	-1.133	0.523	-	-	4.687
2|3	-0.594	0.524	-	-	1.284
3|4	0.243	0.531	-	-	0.209
0|1 Group	-0.803 (n.s.)	0.601	-	-	1.787
1|2 Group	0.229 (n.s.)	0.522	-	-	0.192
2|3 Group	0.419 (n.s.)	0.546	-	-	0.588
3|4 Group	0.706 (n.s.)	0.627	-	-	1.271
Variant (b) – Sequential contrast coding					
pre-MACRO	0.580 (n.s)	0.435	1.786 [0.845; ∞]	-	1.623
pre-BONDS	0.615 (n.s.)	0.527	1.849 [0.721; ∞]	-	1.153
Grade	**-0.428 *****	**0.152**	**0.652 [0; 0.838]**	**-**	**7.885**
pre-MARCO x Group	-0.523 (n.s.)	0.832	0.592 [0; 2.330]	-	0.396
pre-BONDS x Group	**1.828 ***	**1.000**	**6.224 [1.200; ∞]**	**-**	**3.338**
Threshold Coefficients					
0|1	-2.524	0.476	-	-	28.111
1|2	-1.725	0.451	-	-	14.608

View Article Online
DOI: 10.1039/D6RP00153J

2|3	-1.186	0.443	-	-	7.161
3|4	-0.349	0.439	-	-	0.632
0|1 Group	**-1.063 ***	**0.593**	**-**	**-**	**3.222**
1|2 Group	-0.032 (n.s.)	0.435	-	-	0.005
2|3 Group	0.159 (n.s.)	0.420	-	-	0.143
3|4 Group	0.445 (n.s.)	0.462	-	-	0.927

As pre-COLSIMP and pre-COLEXP showed only a few observations and were only present in the KEG subgroup, only pre-MACRO and pre-BONDS were included in the model as pre-sub-modes. The inclusion of SRI did not enhance model fit, leading to its exclusion as an explanatory variable in the final reported model. The variable Group violated a test of the Proportional Odds Assumption (POA) and also demonstrated that the effects at the individual thresholds indeed diverged substantially in a check of effects using a Partial Proportional Odds (PPO) model. Therefore, in this case, the PPO model was retained, in which the POA assumption for Group is relaxed. Accordingly, no global estimate for Group is reported, but rather threshold-specific estimates. These threshold-specific estimates should not be interpreted like global regression coefficients but rather as shifts of the respective threshold. A negative estimate corresponds to a downward shift of the threshold (which, in turn, facilitates the transition from the lower to the higher category). A positive estimate should be interpreted in the opposite way (Christensen, 2019). For example, it is shown that Group = 1 (which corresponds to KEG) significantly facilitates the transition from the lowest category (threshold 0|1). For the remaining thresholds, this effect reverses but does not reach significance in any of the cases. Furthermore, pre-checks of separate regressions for both groups (KEG, KOG) showed that the included pre-sub-modes pre-MARCO and pre-BONDS have substantially different effects in the groups. Therefore, we introduced an interaction term, which – due to the ordinal structure of the predictor pre – results in two interaction estimates. Previous analytical steps indicated that pre-MACRO is particularly unfavorable, while pre-BONDS is particularly favorable for development in KEG. Therefore, we also computed one-sided confidence intervals for the interaction terms at this stage. Regarding these latter, for example the odds ratio for pre-BONDS x Group indicates a positive effect in favor of Group = 1 (KEG). However, the 95% confidence interval is relatively wide, suggesting some uncertainty in the estimate, likely due to a rather small but still sufficient number of observations in this category. Model fit indices were as follows: log-likelihood (LLH) = -207.69, null model log-likelihood (LLH_{Null}) = -226.36, and likelihood-ratio test G^2 = 37.33 ($p < .001$). Nagelkerke's pseudo-R^2 was 0.236. Significance levels: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$.

DOI: 10.1039/D6RP00153J

MP Esterification

Table A12: Estimated coefficients and significance levels for the ordinal regression model for the progress of the esterification reaction (MP Esterification).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
pre-MACRO	0.609 (n.s.)	0.602	1.839 [0.683; ∞]	-	1.022
pre-BONDS	0.032 (n.s.)	0.584	1.033 [0.395; ∞]	-	0.003
SRI	**0.388 ***	**0.205**	**1.473 [1.051; ∞]**	**-**	**3.561**
Grade	-0.239 (n.s.)	0.163	0.787 [0; 1.031]	-	2.134
pre-MACRO x Group	**-2.770 ***	**1.551**	**0.063 [0; 0.804]**	**-**	**3.186**
pre-BONDS x Group	1.130 (n.s.)	0.836	3.097 [0.783; ∞]	-	1.831
Threshold Coefficients					
0|1	-1.147	0.475	-	-	5.827
1|2	0.026	0.470	-	-	0.003
2|3	1.058	0.504	-	-	4.418
3|4	2.374	0.626	-	-	14.379
0|1 Group	-0.808 (n.s.)	0.507	-	-	2.538
1|2 Group	0.637 (n.s.)	0.476	-	-	1.790
2|3 Group	0.003 (n.s.)	0.520	-	-	0.0004
3|4 Group	0.038 (n.s.)	0.742	-	-	0.003
Variant (b) – Sequential contrast coding					
pre-MACRO	0.609 (n.s.)	0.602	1.839 [0.683; ∞]	-	1.022
pre-BONDS	-0.578 (n.s.)	0.791	-	0.562 [0.119; 2.648]	0.531
SRI	**0.388 ***	**0.205**	**1.473 [1.051; ∞]**	**-**	**3.561**
Grade	-0.239 (n.s.)	0.163	0.787 [∞; 1.031]	-	2.134
pre-MACRO x Group	**-2.770 ***	**1.551**	**0.063 [∞; 0.804]**	**-**	**3.186**
pre-BONDS x Group	**3.901**	**1.703**	**49.437 [3.000; ∞]**	**-**	**5.244**
Threshold Coefficients					

View Article Online
DOI: 10.1039/D6RP00153J

0|1	-1.361	0.495	-		7.573
1|2	-0.188	0.487	-		0.150
2|3	0.845	0.517	-	-	2.667
3|4	2.160	0.633	-	-	11.635
0|1 Group	-0.262	0.647	-	-	0.163
1|2 Group	**1.184 ***	**0.643**	**-**	**-**	**3.389**
2|3 Group	0.550	0.672	-	-	0.667
3|4 Group	0.584	0.849	-	-	0.446

As pre-COLSIMP and pre-COLEXP showed only a few observations and were only present in the KEG subgroup, only pre-MACRO and pre-BONDS were included in the model as pre-sub-modes. The variable Group violated a test of the Proportional Odds Assumption (POA) and also demonstrated that the effects at the individual thresholds indeed diverged substantially in a check of effects using a Partial Proportional Odds (PPO) model. Therefore, in this case, the PPO model was retained, in which the POA assumption for Group is relaxed. Accordingly, no global estimate for Group is reported, but rather threshold-specific estimates. Furthermore, pre-checks of separate regressions for both groups (KEG, KOG) showed that the included pre-sub-modes pre-MARCO and pre-BONDS have substantially different effects in the groups. Therefore, we introduced an interaction term, which – due to the ordinal structure of the predictor pre – results in two interaction estimates. Previous analytical steps indicated that pre-MACRO is particularly unfavorable, while pre-BONDS is particularly favorable for development in KEG. Therefore, we also computed one-sided confidence intervals for the interaction terms at this stage. However, in variant (b), pre-Bonds x Group exhibits an extremely large effect, drive by the combined influence of the more or less “harmful” neighboring pre-sub-mode pre-MARCRO (which seems to be very unfavorable in KEG) and the favoring effect of pre-BONDS. Given the relatively small number of cases in both pre-sub-modes, this effect should nevertheless be interpreted with caution. Model fit indices were as follows: log-likelihood (LLH) = -202.09, null model log-likelihood (LLH_{Null}) = -211.52, and likelihood-ratio test G^2 = 18.86 ($p < .05$). Nagelkerke’s pseudo-R^2 was 0.126. Significance levels: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$.

View Article Online
DOI: 10.1039/D6RP00153J

ME Combustion

Table A13: Estimated coefficients and significance levels for the ordinal regression model for the end of the combustion reaction (ME Combustion).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
Group	**0.653 ***	**0.326**	**1.922 [1.124; ∞]**	**-**	**4.016**
Grade	**-0.300 ***	**0.138**	**0.740 [0; 0.929]**	**-**	**4.722**
Threshold Coefficients					
0|1	-1.530	0.445	-	-	11.842
1|2	-1.318	0.441	-	-	8.842

Neither the introduction of specific prior knowledge in the form of pre-sub-modes nor the SRI improved the model, so neither of these predictors was included in the final model. In this case, comparative contrast coding of the pre-sub-modes as an ordinal predictor is unnecessary, so only variant (a) is reported. Model fit indices were as follows: log-likelihood (LLH) = -106.40, null model log-likelihood (LLH_{Null}) = -110.07, and likelihood-ratio test $G^2 = 7.33$ ($p < .05$). Nagelkerke's pseudo-R^2 was 0.062. Significance levels: *** $p \le .001$, ** $p \le .01$, * $p \le .05$.

View Article Online
DOI: 10.1039/D6RP00153J

ME Esterification

Table A14: Estimated coefficients and significance levels for the ordinal regression model for the end of the esterification reaction (ME Esterification).

Variable	Estimate	Std. Error	OR [95% one-sided CI]	OR [95% CI]	Wald χ^2
Variant (a) – NOCONC as reference category					
Group	**3.020**	**0.480**	**20.499 [9.301; ∞]**	**-**	**39.540**
Grade	**-0.854**	**0.209**	**0.426 [0; 0.601]**	**-**	**16.654**
Threshold Coefficients					
0|1	-1.514	0.495	-	-	9.370
1|2	0.255	0.498	-	-	0.263
2|3	0.586	0.511	-	-	1.316
0|1 pre-LIMIT	**-1.718 ***	**0.801**	**-**	**-**	**4.384**
1|2 pre-LIMIT	-0.381	0.638	-	-	0.356
2|3 pre-LIMIT	0.090	0.727	-	-	0.016

As only one sub-mode of the ordinally structured predictor (pre-LIMIT) is included in the model, variant (a) with NOCONC as the reference category and variant (b) with contrast coding yield identical results. Therefore, we only report variant (a). The inclusion of SRI did not enhance model fit, leading to its exclusion as an explanatory variable in the final reported model. The variable pre-LIMIT violated a test of the Proportional Odds Assumption (POA) and also demonstrated that the effects at the individual thresholds indeed diverged substantially in a check of effects using a Partial Proportional Odds (PPO) model. Therefore, in this case, the PPO model was retained, in which the POA assumption for pre-LIMIT is relaxed. Accordingly, no global estimate for pre-LIMIT is reported, but rather threshold-specific estimates. Model fit indices were as follows: log-likelihood (LLH) = -156.88, null model log-likelihood (LLH_{Null}) = -217.82, and likelihood-ratio test G^2 = 121.90 ($p < .001$). Nagelkerke's pseudo-R^2 was 0.531. Significance levels: *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$.

View Article Online
DOI: 10.1039/D6RP00153J

I. Developmental statements towards a learning progression

(I) Students explain the composition of matter and the structures that build it by describing the interaction of particles and emergent characteristics based on these interactions. Building on this, they explain that these interactions and thus the course of chemical reactions can be measured and influenced.

(II) On the basis of a permanent movement of particles, students describe the start of reactions by referring to random collisions of the reactants.

(III) Students explain how random collisions in the further course of reactions (i.e., progress of a reaction) can lead to bond breaking, which in turn leads to the reorganization of atoms and the formation of new bonds.

(IV) Regarding (II) and (III), students distinguish between different types of particle collisions (effective vs. ineffective) and explain how kinetic energy and environmental or even structural factors may influence the effectiveness of collisions.

(V) Students describe the end of the reaction based on systemic conditions and explain the state of dynamic equilibrium.

(VI) Overall, students explain how the time course of a reaction is influenced by applying knowledge listed under (I) to (V).

View Article Online
DOI: 10.1039/D6RP00153J

J. Integration of macro- and micro-adaptions in a (digital) learning environment

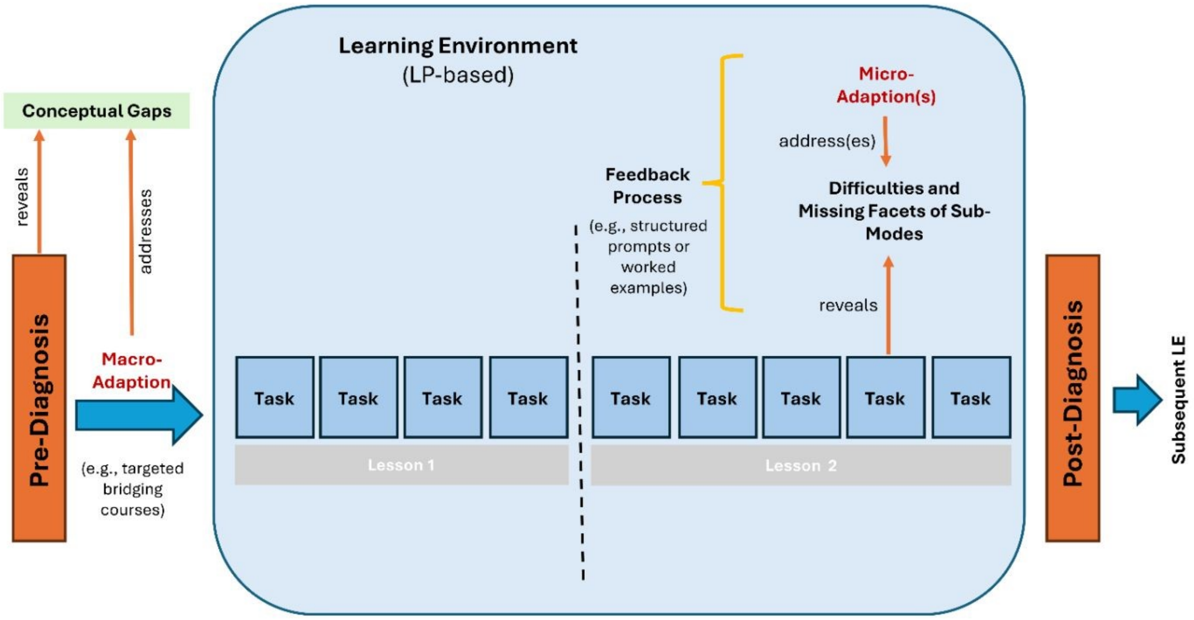

Figure A2: Diagnostic opportunities that can be utilized particularly in digitally supported learning environments, as well as the resulting opportunities for macro- and micro-level adaptations.

View Article Online
DOI: 10.1039/D6RP00153J

K. Methodological literature only cited in the supporting information

Agresti, A., 2002. Categorical Data Analysis, 1st ed, Wiley Series in Probability and Statistics. Wiley. https://doi.org/10.1002/0471249688

Brant, R., 1990. Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. Biometrics 46, 1171. https://doi.org/10.2307/2532457

Fox, J., Monette, G., 1992. Generalized Collinearity Diagnostics. Journal of the American Statistical Association 87, 178–183. https://doi.org/10.1080/01621459.1992.10475190

View Article Online
DOI: 10.1039/D6RP00153J

The data supporting this article have been included as part of the Supplementary Information.