Multi-methods Approach for Domain-Specific Grounding: An ITS for Connection Making in Chemistry

Martina A. Rau and Amanda L. Evenstone

Department of Educational Psychology, University of Wisconsin – Madison
{marau,alevenstone}@wisc.edu

Abstract. Making connections between graphical representations is integral to learning in science, technology, engineering, and mathematical (STEM) fields. However, students often fail to make these connections spontaneously. ITSs are suitable tools to support connection making. Yet, when designing an ITS for connection making, we need to investigate what learning processes and concepts play a role within the specific domain. We describe a multi-methods approach for grounding ITS design in the specific requirements of the target domain. Specifically, we applied this approach to an ITS for connection making in chemistry. We used a theoretical framework that describes potential target learning processes and conducted two empirical studies – using tests, eye tracking, and interviews – to investigate how these learning processes play out in the chemistry domain. We illustrate how our findings inform the design of a chemistry tutor. Initial pilot study results suggest that the ITS promotes learning processes that are productive in chemistry.

Keywords: Connection making, multiple representations, empirically grounded design, multi-methods approach, chemistry.

1 Introduction

The ability to make connections between graphical representations is integral to learning in science, technology, engineering, and mathematical (STEM) fields [1]. For instance, to learn about chemical bonding, students need to make connections between Lewis structures, ball-and-stick figures, space-filling models, and electrostatic potential maps (EPMs; see Figure 1). Connection making is a difficult task that students often do not engage in spontaneously, even though it is critical to their learning [1-2]. Hence, they need support to make these connections [3]. Recent research indicates that intelligent tutoring systems (ITSs) can be effective in supporting connection

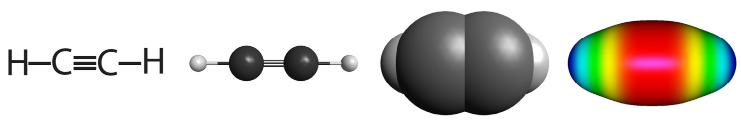

Fig. 1. Graphical representations of ethyne: Lewis, ball-and-stick, space-filling, EPM

S. Trausan-Matu et al. (Eds.): ITS 2014, LNCS 8474, pp. 426–435, 2014.

making [4]. However, to design effective connection making support, we need to investigate which specific learning processes play a role within the target domain. The goal of this paper is to describe a multi-methods approach for grounding the design of an ITS in a particular domain.

Our objective in using this approach is to develop an ITS for connection making that has the potential to significantly enhance students' learning in chemistry. ITS support for connection making is likely to enhance chemistry learning for several reasons. First, the ITS framework of learning through problem solving is in line with the chemistry education literature, which indicates that problem-solving activities can significantly enhance conceptual learning [5], especially when they include graphical representations [6]. Second, even though several educational technologies for chemistry learning exist [7-9], this research is novel because none of them provide explicit and adaptive support for connection making between graphical representations. Finally, the chemistry education literature widely acknowledges that connection making is one of the major stumbling points in chemistry education [10].

In this paper, we describe a new approach to ground the design of this ITS in the chemistry domain. Specifically, we describe how integrating multiple methods provided answers to the following questions: First, which learning processes are important in chemistry and should be supported by the ITS? Second, what problem-solving behaviors should the ITS foster? Third, which chemical bonding concepts should the ITS target? Finally, does the resulting ITS enhance productive learning processes? Even though we address these questions within the chemistry domain, we believe that our approach is a first step towards creating a principled methodology for informing the design of an ITS by the requirements of the specific target domain.

2 Domain-Specific Grounding of Connection-Making Support

2.1 Theoretical Framework

To inform the design of ITS support for connection making, we draw on a theoretical framework, which proposes that two types of connection-making abilities play a role in domain expertise [4]. *Sense-making ability* means that a student can relate aspects that correspond to one another across representations because they depict the same concept (e.g., in the example shown in Fig. 1, relating the local negative charge that results from the triple bond shown in the Lewis structure to the region of high electron density depicted by the red color in the EPM). *Perceptual fluency* is the ability to rapidly and effortlessly find representations that depict the same concept, by relying on perceptual characteristics [11] (e.g., by rapidly seeing that the representations in Fig. 1 show the same molecule based on their linear geometry). The chemistry education literature suggests that both sense-making ability [9, 12] and perceptual fluency in connection making [9-10] are important aspects of chemistry expertise.

We conducted two empirical studies that instantiate this framework for the specific domain of chemistry. Study 1 investigates whether sense-making ability and perceptual fluency are indeed separate connection-making abilities in chemistry. Study 2 investigates the domain-specific aspects of sense-making ability.

2.2 Study 1: Assessment of Sense-Making Ability and Perceptual Fluency

The chemistry education literature documents the importance of both sense-making ability and perceptual fluency in connection making [9]. Confirming the claim that these are indeed distinct abilities is a prerequisite for the design of separate activities to support each of these abilities. To address this question, we conducted an *a priori* factor analysis on an assessment of sense-making ability and perceptual fluency.

Method. Undergraduate and graduate chemistry students with varying levels of expertise were recruited through emails and fliers to take a 30-40 minute online test. 118 students started; 44 students completed the test. We consider resulting missing data to be at random because the item order was at random. The test contained one question about chemistry courses taken, 16 multiple-choice items on sense-making ability (8 on similarities, 8 on differences), and 9 multiple-choice items on perceptual fluency.

Results. We used the SPSS AMOS software to compare three models: a 1-factor model (not distinguishing sense-making ability and perceptual fluency), a 2-factor model (sense-making ability and fluency), and a 3-factor model (sense-making similarities, sense-making differences, and fluency). We excluded missing values (resulting from incomplete tests) on an item-by-item basis. To compare the fit of the tested models, we used root mean squared error (RMSE). The results show that the 3-factor model (RMSE = .072) and the 2-factor model (RMSE = .082) both yielded a better fit than the 1-factor model (RMSE = .088). Because the sense-differences and sense-similarity factors in the 3-factor model correlated highly with $r = .93$, we choose the 2-factor model for further analyses. The resulting two factors, sense-making ability and perceptual fluency, correlate moderately with $r = .62$.

A repeated measures ANOVA showed that students performed significantly better on the sense-making scale ($M = .75$; $SD = .12$) than on the fluency scale ($M = .62$; $SD = .24$; $p < .01$). To investigate the relation of these two abilities with chemistry proficiency, we conducted a regression of the number of chemistry courses taken. The number of courses taken is associated with marginally higher sense-making ability ($\beta = .22$, $p < .10$), and with significantly higher perceptual fluency ($\beta = .448$, $p < .01$).

Discussion. The finding that sense-making ability and perceptual fluency are separate skills in chemistry is in line with the chemistry education literature [9-10, 12] and supports the design of separate activities for these connection-making abilities.

The finding that students have higher sense-making ability than fluency is not surprising: it mimics a current trend in educational practice because most research on connection making focuses solely on sense-making processes [3]. Only recently has perceptual fluency gained attention in the education and psychology literature [11]. Thus, our data encourages the design of an ITS that targets perceptual fluency.

The finding that chemistry proficiency (approximated by the number of courses taken) is a better predictor of perceptual fluency than of sense-making ability is surprising. It seems that chemistry instruction does not sufficiently target the ability to make sense of connections between graphical representations. Given that students' performance on the sense-making scale is far from perfect (*M* = .75; *SD* = .12), there is an instructional need for an intervention that targets students' sense-making ability.

2.3 Study 2: Eye Tracking and Interview Study on Sense-Making Ability

The ability to make sense of the connections between representations involves understanding similarities and differences between different graphical representations. The goal of Study 2 was to investigate the relation between students' ability to identify similarities and differences between representations and their reasoning about domain-relevant concepts. Furthermore, our goal was to identify specific concepts that students struggled with when making connections. Study 2 combined eye-tracking and interview data. This procedure allows us to investigate which visual attention patterns are associated with low and high quality connections.

Method. Twenty-six students participated in Study 2 (21 undergraduate and 5 graduate chemistry students). Sessions took place in the laboratory and lasted 30-45 min. Students were asked to describe similarities and differences between two graphical representations of the same molecule (similar to those in Fig. 1). Students performed this task on an SMI RED250 eye tracker. All verbal responses were audiotaped.

To analyze the eye-tracking data, we created areas of interest (AOIs) for each representation. We considered two measures. First, we considered frequency of switching between AOIs, which is used to indicate perceptual integration [13]. We computed AOI switches as the number of times a fixation on one AOI was followed by another. Second, we considered second-inspection durations. First inspections of an AOI is often considered to indicate initial processing of material that occurs (to a certain extent) automatically [14]. Fixations after the first inspection (i.e., when a student re-inspects an AOI) are considered to reflect intentional processing to integrate the information with other information [14]. We computed the second-inspection durations as the sum of fixation durations that occurred after the initial fixation on a given AOI.

Table 1. First level of the interview coding scheme

Code	Definition (Example)
Surface	Student makes a connection between representations, based on some conceptually irrelevant feature ("um so they're both like red on the top")
Similarities	Student refers to a structural feature of representations that depict the same concept ("the space-filling model and the EPM both in shape are very similar cause they show the electron cloud")
Differences	Student refers to a structural feature of two representations that differs between representations or to information that differs between representations
Inference	Student explains a concept that goes beyond what is depicted ("this [the EPM] just shows that on the oxygen it's more reactive because there's lone

To analyze the interview data, we applied a two-level coding scheme. First-level codes were adapted from prior research on connection making [2]. Specifically, we distinguished connections based on surface features, similarities, or differences, and whether students made inferences about concepts not explicitly shown in the representations. Table 1 provides descriptions and examples for first-level codes. We constructed the second-level codes bottom up: by collecting concepts that were mentioned during the interview and then coding for their occurrence. Interrater reliability was good with 85% agreement for first-level codes and 72.9% for second-level codes.

Results. First, we analyzed how the eye-tracking data relates to the first-level interview codes (see Table 1). Three participants were excluded from the analysis because the eye-tracking ratio was below 85%. A regression of second-inspection durations on first-level codes showed that longer second-inspection durations were associated with significantly more surface connections ($\beta = .60$, $p < .01$), and marginally more differences ($\beta = .39$, $p = .06$). There was no association of second-inspection durations with similarities. A regression of AOI switches on first-level codes showed that more AOI switches were associated with significantly more surface connections ($\beta = .55$, $p < .01$). There was no association with similarities or differences. In turn, a regression of surface connections, similarities, and differences on inferences showed that difference utterances were associated with significantly more inferences ($\beta = .51$, $p < .01$). There were no associations between similarity or surface utterances and inferences.

Next, we analyzed the second-level interview codes. We identified concepts related to the topics of atom identity (symbol, number of electrons, CPK color coding, general identity information), molecule structure (bond angle, bond length, conformation, geometry, atomic radii, electron cloud), energy (steric interactions, relative energy), electrons (core, valence, shared, lone), atomic structure (shells, orbitals, hybridization potential, spin states), and bonding (type, electronegativity, charge distribution). To get insights into which concepts are particularly difficult for undergraduates, we compared the relative frequency of a concept being discussed by graduate versus undergraduate students. We used differences larger than 1 SD to indicate that undergraduates were unlikely to point out this difference, even though it relates to an important concept. We found that the most difficult concepts for undergraduates were CPK color coding, bond angle, atomic radii, relative energy, bonding type, and reactivity. In addition, undergraduate students were less likely use these concepts to make inferences about the behavior of electrons, atoms, and molecules to explain bonding.

Discussion. Our findings show no clear positive effects of commonly used measures of visual attention. Integrating the eye-tracking data with first-level interview codes allowed us to disambiguate the effects of eye-tracking measures on students' reasoning about domain-relevant concepts. Students who switched more frequently between representations were more likely to focus on surface-level connections. Students with longer second-inspection durations were more likely to notice surface features and differences between representations. Only difference-connections were associated with making more inferences about domain-relevant concepts.

It is surprising that we found no positive associations between noticing similarities between representations and making inferences about chemistry concepts. It may be that expertise in chemistry relies on the use of different graphical representations for complementary purposes, rather than in using them interchangeably because they provide similar information. Indeed, this interpretation aligns with the literature on how chemistry experts use representations [15]. Consequently, we hypothesize that ITS support for connection making in chemistry should focus on how different graphical representations depict *complementary* information, rather than how they depict *similar* concepts. To do so, the ITS should help students to redirect (after initial inspection) their attention to the representations and to focus on them for a longer duration, rather than to frequently switch between different representations.

Furthermore, our findings suggest that the ITS should target the concepts of CPK color coding, bond angle, atomic radii, relative energy, bonding type, and reactivity. These concepts may be more difficult because they are more complex: they are typically used to reason about bonding phenomena that involve the interaction of one molecule with additional atoms and molecules rather than about the structure of individual atoms and molecules.

3 Design of a Chemistry Tutor for Connection Making

Study 1 encourages developing an ITS for chemistry that targets sense-making ability and perceptual fluency through separate activities. Study 2 suggests that sense-making activities should focus on differences between representations, not on similarities. Here we describe how these findings informed the design of a chemistry tutor.

3.1 Tutor Design

In line with prior research [3], *sense-making activities* are designed to help students in relating conceptually relevant aspects of different graphical representations. As Study 2 suggests, we focus on differences between representations in providing complementary information. Sense-making activities involve three parts. Consider a problem that targets one of the concepts that we found to be particularly difficult in Study 2: bonding type and electron behavior (Fig. 2). Students identify the type of bond between atoms and make inferences about how electrons are distributed across the molecule. First, they solve this problem with one representation (e.g., a Lewis structure, see Fig. 2A). Second, they solve a corresponding problem with another representation (e.g., an EPM, see Fig. 2B). Third, students are prompted to explain differences between representations (e.g., the local negative charge is shown by a larger number of electron-dots shown in Lewis structures, and by red color in EPMs; Fig. 2C).

The design of the *fluency-building activities* is based on Kellman and colleagues' perceptual learning paradigm [11]. Rather than focusing on why or how different representations correspond to one another, fluency-building support aims at helping students become faster and more efficient at extracting relevant information from the representations based on repeated experience with a large variety of problems. Thus, the fluency-building activities provide numerous practice opportunities to find corresponding graphical representations based on their perceptual properties. Fig. 3 shows two sample problems in which students have to choose a representation that show the same molecule. Choices are designed to contrast which perceptual aspects provide relevant information. For instance, to solve the example on the left-hand side of Fig. 3, students have to attend to how EPMs depict the geometry of the molecule. To solve the example on the right-hand side, students need to attend to the lone pair in Lewis structures, which have implications for electronegativity that the EPM depicts as color. Students receive a series of these problems and are encouraged to solve them fast, by using perceptual properties and without overthinking the problem.

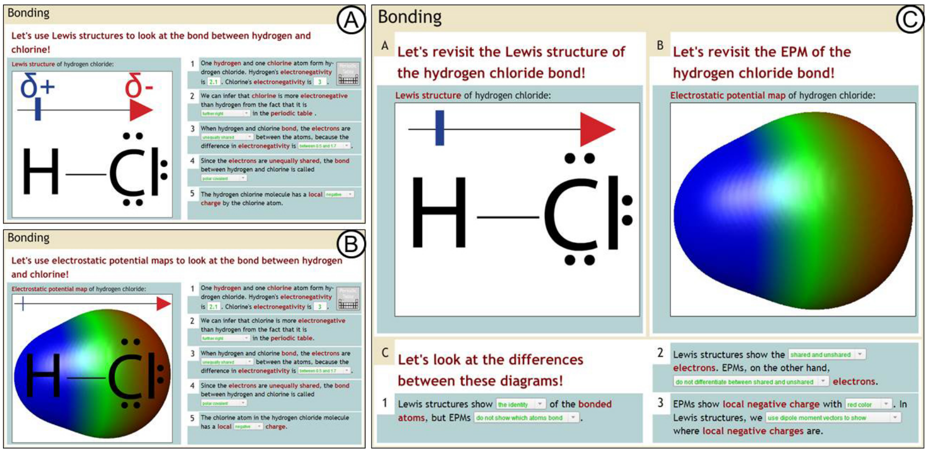

Fig. 2. Sense-making problems

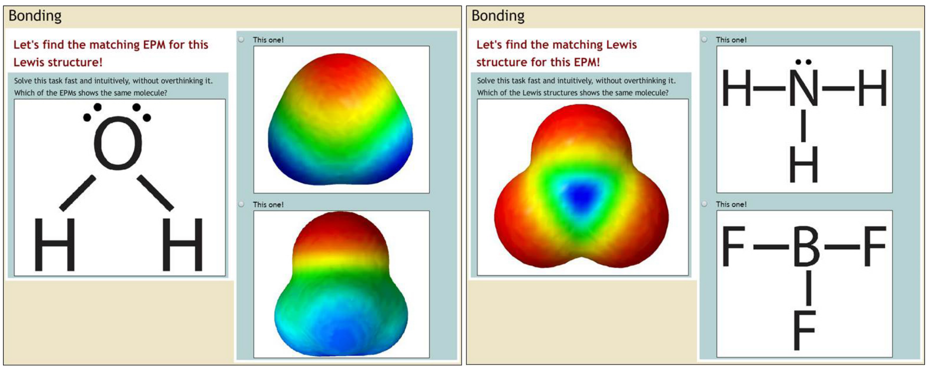

Fig. 3. Fluency-building problems.

3.2 Initial Pilot Results

We collected initial pilot data from four students who worked with a handful of sense-making and fluency-building prototypes. During the pilot sessions, we collected eye-tracking data, interview data, and tutor log data. The interview data suggests that students like the tutor activities because they contain multiple graphical representations. For instance, one student commented, "I think it does a good job at showing multiple layouts instead of just one, so one can understand". The small sample size did not warrant a quantitative analysis of the eye-tracking data. Instead, we viewed the eye-gaze recordings and counted the number of times a student reinspected a graphical representation. For sense-making activities, this qualitative analysis suggests that impasses and reflection prompts (see Fig. 2C) are associated with subsequent reinspection of the representations. For fluency-building problems, we found that students frequently switch between representations. Finally, the log data showed that the reflection prompts (see Fig. 2C) had higher-than-average error rates. Fluency-building activities had a lower average error rate than sense-making problems.

In addition, we collected pre- and post-test data from three students in a second pilot study who worked with a fully-functioning version the ITS for one hour. We found learning gains of 16 percent points on sense-making items, 27 percent points on fluency items, and 7 percent points on transfer items about chemistry concepts.

3.3 Discussion

With respect to the sense-making activities, the pilot log data shows that sense-making prompts are challenging. This observation is in line with the finding in Study 1 that sense-making problems are difficult and further supports the conclusion that we need to support students' sense-making abilities, especially since Study 2 shows that noticing differences between representations is associated with conceptual inferences. Our qualitative analysis of the eye-tracking data suggests that impasses and prompts lead to reinspections of representations. This observation is promising because Study 2 showed that longer second-fixation durations are associated with inferences by helping students notice differences between representations. Thus, the pilot data suggests that sense-making activities enhance productive visual attention behaviors.

With respect to the fluency-building activities, further investigation is needed. The fact that the log data suggests that fluency-building activities are easier than sense-making activities stands in contrast to the finding of Study 1 that students have lower perceptual fluency than sense-making ability. On the one hand, one might conclude that the current design of the fluency-building activities enhances superficial visual processing because they are not difficult enough. On the other hand, we cannot necessarily draw the conclusion that frequent switching between representations and low error rates are associated with low learning gains, because the finding from Study 2 that frequent switching is associated with surface connections was based on an investigation of only sense-making items (not of perceptual fluency items).

Finally, pilot results on pretest to postest learning gains indicates that the ITS is effective in improving students' sense-making ability, perceptual fluency, and transfer of conceptual knowledge. An experiment testing the effectiveness of the sense-making and fluency-building components of the ITS is currently under way. Specifically, we will analyze mediation effects of eye-gaze behaviors, conceptual reasoning, and problem-solving behaviors on students' pretest to posttest learning gains.

4 Conclusion and Future Work

We described a multi-methods approach to ground the design of an ITS in the requirements specific to the target domain. Our goal in applying this approach to the chemistry domain was to inform the design of an ITS for connection making. Our empirical approach built on a theoretical framework that proposed two separate abilities: sense-making ability and perceptual fluency. We then conducted an assessment study that supports the existence of these two connection-making abilities in the chemistry domain. Even though this finding is in line with the chemistry education literature, which states that both skills are important aspects of chemistry expertise [9], our study is (to the best of our knowledge) the first to provide empirical support for this claim. Next, we conducted a study that combined eye-tracking and interview data to investigate which learning processes and concepts are most important with respect to sense-making ability. We found that making sense of differences between representations is more important than making sense of similarities between representations. Our data suggests that the visual mechanism by which students attend to differences between representations is to reinspect graphical representations rather than to frequently switch between representations (possibly among others). Furthermore, we identified several aspects of representations that undergraduates fail to identify spontaneously even though they constitute important chemistry concepts. Finally, our initial pilot results indicate that the ITS design enhances productive learning processes, that students perceive it as valuable, and that it leads to learning gains.

A limitation of the research described in the present paper is that our data are correlational in nature, but not causal. The results from Study 1 lead to the *prediction* that providing separate activities to support sense-making ability and perceptual fluency enhances students' learning in chemistry. Furthermore, the findings from Study 2 lead to the *prediction* that sense-making activities will enhance students' learning if they emphasize differences between representations rather than similarities, and if they help students to visually reinspect representations. The next step in our research is to experimentally test these predictions. We are currently conducting an experiment to evaluate the effectiveness of sense-making and fluency-building activities based on pretest to posttest learning gains, and to contrast whether (as hypothesized) students learn best when working with both sense-making and fluency-building activities, compared to working with either type of activity alone. Furthermore, we use the eye-tracking and interview measures described above to analyze whether (and how) students' visual attention patterns and connection-making utterances mediate the anticipated effects of the sense-making and fluency-building activities.

In sum, by using a multi-methods approach to ground ITS design in the specific requirements of the chemistry domain, we developed a system that appears to enhance productive learning processes and that addresses educational needs. Furthermore, this approach equips us with an initial theoretical model of how students' connection making might enhance their learning in chemistry and with a set of eye-tracking and interview measures that we can use to evaluate the effectiveness of the ITS. We conclude that our approach presents a useful methodology to identify *domain-specific* aspects that should shape the design of ITSs with multiple graphical representations.

Acknowledgements. This work was supported by the UW Madison Graduate School and the Wisconsin Center for Education Research. We thank Joe Michaelis, Abigail Dreps, Brady Cleveland, William Keesler, Taryn Gordon, and Theresa Shim for their contributions.

References

1. Ainsworth, S.: DeFT: A conceptual framework for considering learning with multiple representations. Learning and Instruction 16, 183–198 (2006)
2. Rau, M.A., Rummel, N., Aleven, V., Pacilio, L., Tunc-Pekkan, Z.: How to schedule multiple graphical representations? A classroom experiment with an intelligent tutoring system for fractions. In: Aalst, J.V., Thompson, K., Jacobson, M.J., Reimann, P. (eds.) Proceedings of the 10th International Conference of the learning Sciences (ICLS 2012) vol. 1, pp. 64-71. ISLS, Sydney, Australia (2012)
3. Bodemer, D., Faust, U.: External and mental referencing of multiple representations. Computers in Human Behavior 22, 27–42 (2006)
4. Rau, M.A., Aleven, V., Rummel, N., Rohrbach, S.: Sense Making Alone Doesn't Do It: Fluency Matters Too! ITS Support for Robust Learning with Multiple Representations. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 174–184. Springer, Heidelberg (2012)
5. Bodner, G.M., Domin, D.S.: Mental models: The role of representations in problem solving in chemistry. University Chemistry Education 4, 24–30 (2000)
6. Bowen, C.W.: Representational systems used by graduate students while problem solving in organic synthesis. Journal of Research in Science Teaching 27, 351–370 (1990)
7. Kozma, R., Russell, J.: Multimedia learning of chemistry. Cambridge handbook of multimedia learning, pp. 409-428 (2005)
8. Stieff, M.: Connected chemistry-A novel modeling environment for the chemistry classroom. Journal of Chemical Education 82, 489–493 (2005)
9. Wu, H.K., Krajcik, J.S., Soloway, E.: Promoting understanding of chemical representations: Students' use of a visualization tool in the classroom. Journal of research in science teaching 38, 821–842 (2001)
10. Gilbert, J.K., Treagust, D.F.: Towards a Coherent Model for Macro, Submicro and Symbolic Representations in Chemical Education. In: Gilbert, J.K., Treagust, D.F. (eds.) Multiple Representations in Chemical Education, pp. 333–350. Springer, Netherlands (2009)
11. Kellman, P.J., Massey, C.M., Son, J.Y.: Perceptual Learning Modules in Mathematics: Enhancing Students' Pattern Recognition, Structure Extraction, and Fluency. Topics in Cognitive Science 2, 285–305 (2009)
12. Cheng, M., Gilbert, J.K.: Towards a better utilization of diagrams in research into the use of representative levels in chemical education. In: Gilbert, J.K., Treagust, D.F. (eds.) Multiple Representations in Chemical Education, pp. 191–208. Springer, Netherlands (2009)
13. Johnson, C.I., Mayer, R.E.: An eye movement analysis of the spatial contiguity effect in multimedia learning. Journal of Experimental Psychology: Applied 18, 178–191 (2012)
14. Mason, L., Pluchino, P., Tornatora, M.C.: Effects of Picture Labeling on Science Text Processing and Learning: Evidence From Eye Movements. Reading Research Quarterly 48, 199–214 (2013)
15. Kozma, R., Russell, J.: Students becoming chemists: Developing representationl competence. In: Visualization in Science Education, pp. 121–145. Springer, Netherlands (2005)