HOW EXPERTS AND NOVICES NAVIGATE CHEMISTRY REPRESENTATIONS – AN EYE-TRACKING INVESTIGATION

Prajakt Pande[1], Prateek Shah[2] & Sanjay Chandrasekharan[1]

[1]Homi Bhabha Centre for Science Education, TIFR, Mumbai (India),
[2]Indian Institute of Management, Ahmedabad (India)

prajaktp@hbcse.tifr.res.in, prateekshah@iimahd.ernet.in, sanjay@hbcse.tifr.res.in

Experts are known to form more sophisticated conceptual associations between multiple external representations (MERs) than novices, however, the cognitive mechanisms underlying this ability in chemistry is not well understood. We attempt to characterize expert-novice differences in terms of the way they mentally process chemistry MERs. In this study, chemistry professors (experts) & undergrads (novices) view & categorize MERs. Using eye-tracking, we capture fine-grained data about participants' gaze patterns while they view given MERs, which we then correlate with the quality of categories they generate as well as justifications they provide for those categories. The professors tend to form chemically meaningful relationships between MERs than do undergrads. Eye-tracking data reveal differences between the two groups, in navigating chemical equations.

INTRODUCTION

Chemistry deals with complex systems, entities & phenomena that often cannot be directly perceived (e.g. atoms, chemical reactions, etc.) These imperceptible systems are understood at multiple levels of detail (electronic configuration, stereo-chemistry, stoichiometric ratios etc.), using multiple external representations (MERs), such as reaction mechanisms, molecular diagrams, graphs & equations, at each level. The ability to generate & use these MERs in an integrated fashion (for conceptualization, discovery & communication) is indicative of expertise in chemistry. This skill-set is collectively known as representational competence (abbreviated as RC, Kozma & Russell, 1997). Developing RC (expertise over MERs) is an important goal of chemistry education. Problems & difficulties in teaching/learning chemistry are attributed to difficulties in understanding the MERs in chemistry (Johnstone, 1991 & 1993; Kozma & Russell, 1997; Gilbert & Treagust, 2009).

A significant strand of research in chemistry education reports descriptions of students' use of multiple representations, transformations of these representations, and the difficulties students face while doing both of the above. Studies show that students fail to associate the symbols and numbers with substances and phenomena (in other words relate MERs and the information they convey; Herron & Greenbowe, 1986; Nurrenbern & Pickering, 1987; Sanger & Phelps, 2007), primarily due to a lack of clarity on basic concepts such as oxidation numbers, ionic charge, atoms and atomic structure, formal rules for writing molecular formulae, as well as meaning of subscript numbers and brackets and coefficients (Savoy, 1988). Ben-Zvi, Eylon and Silberstein, (1988) propose that students' thinking about phenomena relies primarily on perceptual/sensory information but since current pedagogical practices hardly provide perceptual/sensory assistance, students do not understand chemical symbols in terms of their macro and micro-level instantiations. Johnstone's model of three thinking levels (Johnstone, 1982) and versions thereof, describe three different levels of chemistry MERs: (a) macro level, where one sees and handles materials, observes and describes phenomena and their properties, such as color, flammability, solubility, (b) symbolic level, where one represents chemical substances and phenomena using symbols, formulas, equations and conventions, and (c) submicro level, at which one explains the nature of chemical substances, mechanisms of reactions, and the underlying molecular/atomic interactions. Johnstone (1991) attributes students' difficulties in learning chemistry to the difficulty in simultaneously handling MERs distributed across these three levels as a result of the limited capacity of the human working memory (Ben-Zvi, Eylon & Silberstein, 1988; Justi & Gilbert, 2002; Kozma & Russell, 1997; Mayer, 2002; Sirhan, 2007).

Another strand of research attempts to characterize and examine RC, and describes expert-novice differences in terms of use of MERs. For instance, researchers demonstrate using eye-tracking, that students mainly concentrate on graphical and model representations in animations and often ignore equations, when interacting with a multi-representational molecular mechanics animation (Stieff, Hegarty & Deslongchamps, 2011). While students face difficulties in producing static representations (e.g. sketches; Madden, Jones & Rahm, 2011) of the (imagined) dynamic particulate interactions, experts, on the other hand, seem to better transform between static (such as equation & graphs) and dynamic representations (such as reaction mechanisms; Wu & Shah, 2004; Nakhleh & Postek, 2008). Kozma and Russell (2005), identify specific skills among chemistry experts, viz., (a) using representations to describe chemical phenomena, (b) generating and/or selecting appropriate MERs according to specific needs, (c) identifying and analyzing different features of MERs, (d) comparing and contrasting different MERs, (e) making connections across different representations, relating/mapping features between MERs, (f) understanding that the MERs correspond to phenomena but are distinct from them, and (g) using MERs to support claims, draw inferences, and make predictions. Levy and Wilensky (2009) suggest that understanding chemical phenomena involves building of internal (mental) models that simulate the behaviors of many individual molecules/atoms, their collective behaviors and properties, and effects of various parameters on such behaviors.

Current characterizations of student difficulties and/or RC in chemistry can summarily be categorized into – cognitive load based explanations (expert is better able to handle the cognitive load by employing cognitive strategies such as information chunking, whereas novices lack such skills, Cook, 2006; Johnstone, 1982), context & practice based accounts (students lack exposure to these while experts have had ample exposure, Ben-Zvi, Eylon & Silberstein, 1988; Nelson, 2002; Tsaparlis, 2009), and conceptual understanding/prior-knowledge based explanations (which say that students have superficial understanding and low prior knowledge making it difficult for them to understand MERs; Cook, 2006; Nitz & Tippett, 2012). Ultimately, all these accounts boil down to the classical information processing framework emphasizing cognitive load and strategies to lower/handle it. Such accounts do not seek to provide a detailed understanding of the cognitive mechanisms underlying the processing of MERs, and thus offer only a rather superficial account of MER integration.

Our research attempts to characterize RC by developing models of the cognitive mechanisms underlying the processing of MERs, particularly integration of MERs (which is how we define RC), and suggest design principles for interventions. In this study, chemistry professors (experts) & undergrads (novices) view & categorize MERs. Using eye-tracking, we capture fine-grained data about participants' gaze patterns while they view given MERs, which we then correlate with the quality of categories they generate as well as justifications they provide for those categories.

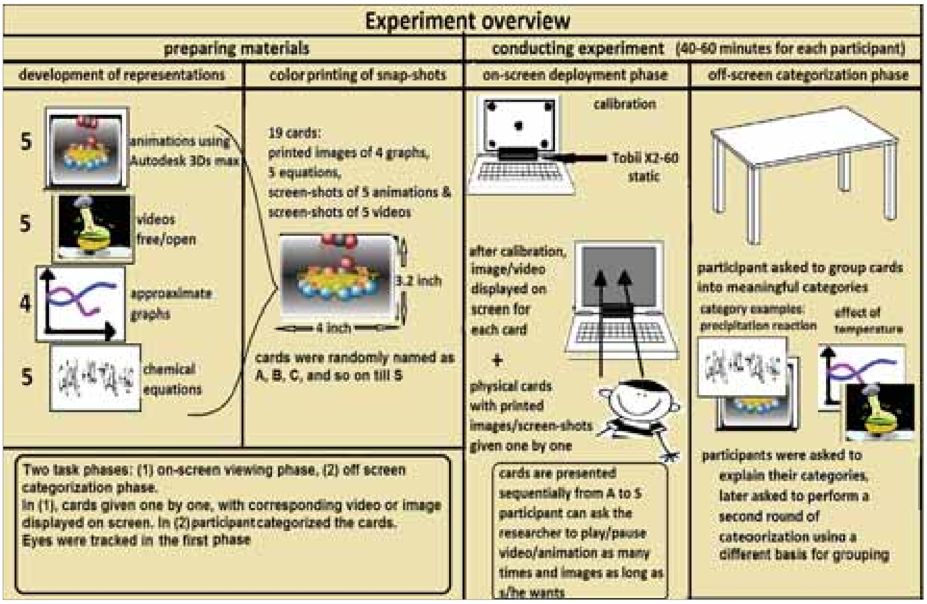
We used Tobii X2-60 static eye-tracker to capture fine-grained data on student eye-movement and gaze patterns across MERs presented to (and handled by) them. Our preliminary analysis confirms earlier reports on novices' surface-feature-based exploration of MERs, but adds details of eye-gaze patterns.

EXPERIMENTAL SETUP

An MER categorization task (from Kozma & Russell, 1997) was conducted with six chemistry undergrad students (3 girls). We describe below the two phases of the study.

Preparing Task Material

Materials for the categorization experiment included different representations for five pre-determined general chemical reactions. There were four representations corresponding to each reaction – a chemical equation, a graph (except for the precipitation reaction), a video of laboratory personnel performing the reaction in a laboratory, and a bare 3D molecular animation (that depicted only the reaction mechanism at molecular level). We developed bare 3D molecular animations for the five chemical reactions. Each animation depicts only the molecular dynamics of that reaction, and does not have any other embedded representations, such as text, narrative, graphs or equations; thus, only one kind of representation. Free videos of the five chemical reactions (being performed in laboratories) from on-line sources were used. Chemical equations and approximate graphs for each reaction (except for the precipitation reaction that had no graph) were generated. This resulted in 19 representations corresponding to five different chemical phenomena. To make these representations more convenient for physical handling, the image of each representation (for animation and video, snapshot of an important moment as an image) was color printed and pasted on a 3x4 inch cardboard, generating 19 cards. Figure 1 depicts preparation and execution of the experiment in detail.



19 cards:
printed images of 4 graphs,
5 equations,
screen-shots of 5 animations &
screen-shots of 5 videos

cards were randomly named as A, B, C, and so on till S

after calibration, image/video displayed on screen for each card + physical cards with printed images/screen-shots given one by one

participant asked to group cards into meaningful categories

Two task phases: (1) on-screen viewing phase, (2) off screen categorization phase.
In (1), cards given one by one, with corresponding video or image displayed on screen. In (2) participant categorized the cards.
Eyes were tracked in the first phase

cards are presented sequentially from A to S participant can ask the researcher to play/pause video/animation as many times and images as long as s/he wants

participants were asked to explain their categories, later asked to perform a second round of cateoorization usino a different basis for grouping

Figure 1: Material development and experimental design details

Running the Experiment

Six chemistry undergrads (3 females) as novices and seven chemistry faculty (4 females) as experts from different university colleges in the city of Mumbai participated in the categorization experiment. Each participant performed the experiment individually. The experiment had two phases:

On-screen phase

Participant was given each of the 19 cards (one after the other, in a pre-determined random order maintained for all participants), and was shown the corresponding image/video on a laptop screen. The participant could observe the images as long, and videos/animation as many times as he/she wanted. Going back to a previously shown representation was not allowed.

Off-screen phase

Once the participant viewed all the 19 representations and had all the cards, he/she was asked to group the cards into meaningful categories. There was no time limit to this phase. They were also asked to explain the different categories made and the basis of categorization (relationship between the cards/representations). The researcher then asked the participant to perform another round of categorization using a different grouping scheme, and explain the grouping criteria.

Data Collection

We used eye-tracking (Tobii X2-60, a static eye-tracker) during the on-screen phase of the task, to obtain fine-grained data about participants' eye-movements and gazes when they viewed the representations (See Pande & Chandrasekharan, 2014, for details eye-tracker setups).

Sources of data collection: (a) for on-screen phase – dynamic eye-movement and fixation data superimposed on the screen-capture video, (b) for off-screen phase – categories made by the participants, their verbal justifications, and side-view video recording of the categorization and justification sessions. The entire session ranged from 40-60 minutes for each participant.

RESEARCH QUESTIONS

1. Do experts make more chemically meaningful associations between MERs than novices?
2. Are there any gaze-pattern differences between experts & novices over static representations? If yes, what differences? How are they related to categorization?

HYPOTHESES

1. In graphs, the total fixation duration for experts would be higher for curves than the axes, as the shape of the curves conveys dynamic information about the phenomena (e.g. sigmoid behavior with time). Instead, novices are likely to spend more time on the axes than curves, as they might find numerical information more relevant to adhere to.
2. On equations, experts' total fixation duration would be more distributed across reactants, arrow, and products, as they would systematically look through each part of the equation and transit between the sub-scripts, super-scripts and coefficients.

Novices would either move randomly or tend to focus either on reactants or products more.

3. Experts would make more long-distance transitions over different parts of the equations, than novices, who would tend to move between closely located elements.

FINDINGS

We report preliminary results on (a) the nature of categories experts and novices make in the first trial of categorization, and the justifications they provide for those categories, (b) statistical analysis of the fixation data on static representations (graphs and equations), and (c) fine-grained process data on how the two groups differ in the way they navigate chemical equations during the viewing phase.

Nature of Categories

We coded the categories of representations participants generated, based on the chemical meaningfulness of relations/connections participants established between different representations, into following five types. (1) Conceptual categories: Chemically meaningful combinations of cards supplemented with correct conceptual description of grouping criteria (e.g. associations of cards depicting equilibrium phenomena, precipitation reaction). (2) Mixed categories: Categories with correct/plausible combinations of cards, with some associations and/or representations explained using chemical concepts while others explained using visual features (e.g. a category made with, say 4 cards depicting equilibrium reaction, of which two cards are explained using the concept of equilibrium while the other two explained based on similarity in features such as heating, or temperature-concentration axes of a graph). (3) Categories based on similarity in visual-features between the representations: Associations of cards explained purely on the basis of visual features of the representations grouped together (e.g. animation showing settling of molecules and a laboratory demonstration exhibiting precipitation; association explained in words such as, ‘both settling down’.) (4) Media-based categories: Complete media-based combinations of cards (e.g. all molecular animations/simulations as a category, all graphs as another, etc.), and (5) Non-sense categories: Incorrect or meaningless combinations of cards not employing falling under any of the above category types (e.g. an association between a precipitation reaction equation with a video showing effect of temperature on a chemical equilibrium.

Experts tend to form more number of mixed as well as conceptual (chemically meaningful) categories than do novices, who tend to associate MERs more often based on their visual features and their medium of representation. This confirms the results from a previous study by Kozma and Russell (1997). The two groups do not seem to differ from each other in terms of the number of non-sense and media-based categories they made. Figure 2 depicts the mean percentage for each type of category generated by experts and novices, during the first round of categorization. A similar trend is observed over second round of categorization.

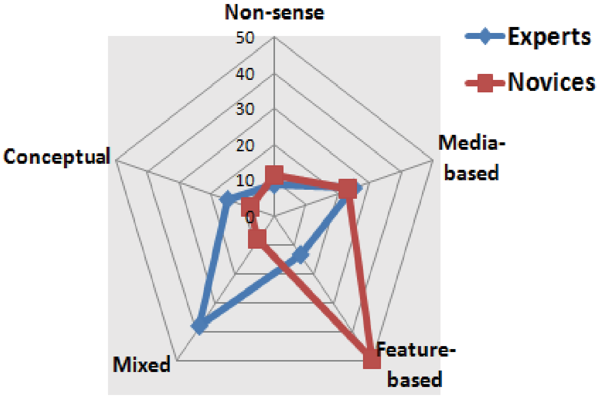


Figure 2: Distribution of participants’ categories across different types

Fixation/Visit Duration Analysis

Fixation duration is a useful statistic to understand the total time spent by a participant viewing a given area of interest (AOI) or part of the representation while viewing it. We found no expected differences between experts and novices. They seem to spend their time viewing the different AOIs roughly similarly, thus rejecting hypotheses (i) and (ii). Both the groups seem to fixate slightly longer on the axes in the graphs, and reactants in the equations.

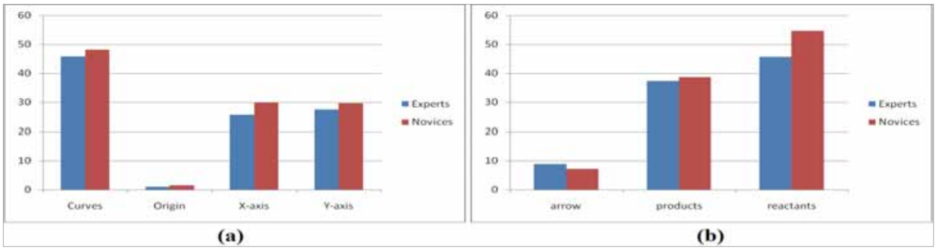


Figure 3: (a) Percent fixation duration on different parts (areas of interest - AOIs) across all four graphs presented, (b) Percent fixation duration on different AOIs across all the five equations.

Since nothing conclusive can be said through the fixation duration statistics, we decided to delve further into the viewing/thought process data. Below we report one aspect of such qualitative data – nature of fixation transitions (jumps).

Nature of Gaze Transitions

Here we report transition data only for equations. We characterized two kinds of transitions viz. long jumps (gaze transitions occurring within two distantly situated AOIs in the space) and short jumps (gaze transitions happening over two closely situated AOIs). For instance, in figure 4, any direct transition between the two reactants (R1 and R2) or between the two products (P1 and P2) would be counted as short jumps, whereas, transitions between the reactant and the product side would be long jumps.

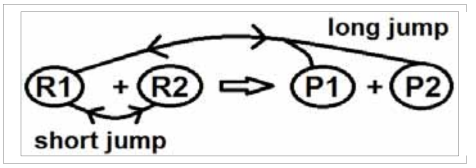


Figure 4: Long and short jumps

Experts performed more number of long jumps than novices on an average, while novices tended to perform more number of shorter jumps than longer jumps in comparison to the experts (results can be considered as partially significant at $p = 0.05$, as the extreme deviations from both groups overlap slightly, apparent in the box plots in figure 5).

Figure 6 depicts a normalized distribution of long jumps performed by experts and novices across all the equations. Experts make significantly higher number of longer jumps than novices. Conversely, they make significantly less number of short jumps than the novices.

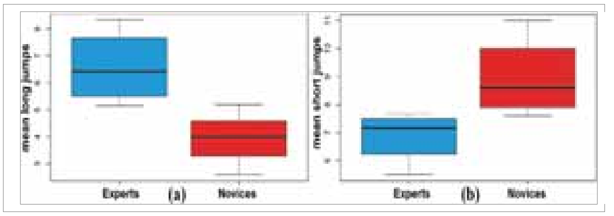


Figure 5: Box plots capturing (a) mean number of long jumps across all equations, (b) mean number of short jumps across all equations.

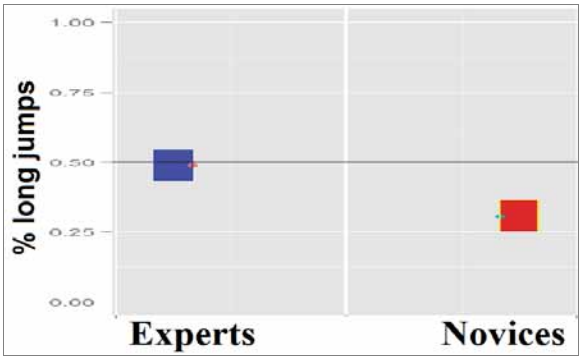


Figure 6: Percentage long jumps performed by experts and novices across all the equations, an inverse would be percent short jumps performed.

CONCLUSION

Our findings confirmed some results from previous literature, and added further details about how experts and novices move their eyes as they navigate (through) the MERs. Experts tend to make chemically meaningful as well as mixed groups of MERs in the categorization task more often than do novices, who tend to relate MERs based on their surface features. The eye tracking data suggests that RC and expertise can be characterized in terms of eye movements and gaze patterns across MERs (behavioral/cognitive markers). Significant differences between experts and novices in the proportion of gaze transitions between distant AOIs (reactants and products) suggest differences in the way they understand the dynamic relationship between reaction components. Experts may be said to imagine the reaction dynamics better, by relating elements between reactants and products, and understanding the points of chemical/substance-level transformations. Novices, on the other hand, seem to look at the chemical equation more linearly.

Further analysis is required to (i) isolate eye movement and navigation patterns related to RC, as well as (ii) comment specifically on the nature of internal/mental representation.

References

Ben-Zvi, R., Eylon, B., & Silberstein, J. (1988). Theories, principles and laws. *Education in Chemistry*, *25*(3), 89-92.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*(6), 1073-1091.

Gilbert, J. K., & Treagust, D. (2009). Towards a coherent model for macro, submicro and symbolic representations in chemical education. In J. K. Gilbert & D. Treagust (Eds.), *Multiple representations in chemical education* (pp. 333–350). The Netherlands: Springer.

Herron, J. D., & Greenbowe, T. J. (1986). What can we do about sue: A case study of competence. *Journal of Chemical Education*, *63*(6), 528.

Johnstone, A. H. (1982). Macro and microchemistry. *School Science Review*, *64*(227), 377-379.

Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of Computer Assisted Learning*, *7*(2), 75-83.

Johnstone, A. H. (1993). The development of chemistry teaching: A changing response to changing demand. *Journal of Chemical Education, 70*(9), 701–705.

Justi R. S., & Gilbert, J. K. (2002). Models and modelling in chemical education. In J. Gilbert, O. De Jong, R. Justi, D. F. Treagust and J. H. Van Driel (Eds.), *Chemical education: Towards research-based practice* (pp. 213-234). Dordrecht: Kluwer Academic Publishers.

Kozma, R. B., & Russell, J. (1997). Multimedia and understanding: Expert and novice responses to different representations of chemical phenomena. *Journal of Research in Science Teaching*, *34*(9), 949-968.

Kozma, R., & Russell, J. (2005). Students becoming chemists: Developing representational competence. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 121-145). Dordrecht: Springer.

Levy, S. T., & Wilensky, U. (2009). Crossing levels and representations: The connected chemistry (CC1) curriculum. *Journal of Science Education and Technology*, *18*(3), 224–242.

Madden, S. P., Jones, L. L., & Rahm, J. (2011). The role of multiple representations in the understanding of ideal gas problems. *Chemistry Education Research & Practice.*, *12*(3), 283-293.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85-139.

Nakhleh, M. B., & Postek B. (2008). Learning chemistry using multiple external representations. In J. K. Gilbert et al., (Eds.), *Visualization: Theory and practice in science education*, (pp. 209–231). NY: Springer.

Nelson, P. G. (2002). Teaching chemistry progressively: From substances, to atoms and molecules, to electrons and nuclei. *Chemistry Education Research & Practice*, *3*(2), 215-228.

Nitz, S., & Tippett, C. D. (2012). Measuring representational competence in science. In E. de Vries & K. Scheiter (Eds.), *Proceedings EARLI special interest group text and graphics: Staging knowledge and* experience: How to take an advantage of representational technologies in education and training (pp. 163-165). Grenoble, France: Université Pierre-Mendès-France.

Nurrenbern, S. C., & Pickering, M. (1987) Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education, 64*(6), 508.

Pande, P., & Chandrasekharan, S. (2014) Eye-tracking in STEM education research: limitations, experiences and possible extensions, In Kinshuk & Murthy, S. (Eds.), *Proceedings of the 6th IEEE International Conference on Technology for Education* (pp. 116-119), Amrita University, Kerala: IEEE Conference Publications/ Conference Publishing Services.

Sanger, M. J., & Phelps, A. J. (2007). What are students thinking when they pick their answer? A content analysis of students' explanations of gas properties. *Journal of Chemical Education, 84*(5), 870–874.

Savoy, L. G. (1988). Balancing chemical equations. *School Science Review*, *69*(249), 713-720.

Sirhan, G. (2007). Learning difficulties in chemistry: An overview. *Journal of Turkish Science Education, 4*(2), 2-20.

Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction*, *29*(1), 123-145.

Tsaparlis, G. (2009). Learning at the macro level: The role of practical work. In J. K. Gilbert & D. Treagust (Eds.), *Multiple representations in chemical education* (pp. 109-136). Dordrecht: Springer.

Wu, H. K., & Shah, P. (2004). Exploring visuospatial thinking in chemistry learning. *Science Education*, *88*(3), 465–492.