Part 2 / Strand 2

Learning Science: Cognitive, Affective, and Social Aspects

Editors: *Florence Le Hebel & Vanessa Kind*

Part 2. Learning Science: Cognitive, Affective, and Social Aspects

Cognitive, affective, and social dimensions in learning science. Design of in-school and extra school learning environments and study of teaching/learning processes. Representational languages and knowledge organisation. Collaborative construction of knowledge.

ESERA 2021
Fostering scientific citizenship in an uncertain world
30 Aug - 3 Sep 2021
Organised by University of Minho, Braga, Portugal

HOW CONSTRUCTED EPISTEMIC EMOTIONS FACILITED AN ELEMENTARY STUDENT'S PARTICIPATION PATTERNS IN SMALL-GROUP SCIENCTIFIC MODELING

Sally B. Gutierezr[1], *Moonhyun Han*[2]

[1]Far Eastern University, Manila, Philippines

[2]Seoul National University of Education, Seoul, South Korea

In this qualitative case study, we purposively selected a Korean female elementary student, anonymously named Susan, and explored which epistemic emotions she constructed during her participation in small-group scientific modeling of the human respiratory system. With six lessons, students in small groups discussed and formulated arguments in order to construct a scientific model that will make them understand the organ structure, mechanisms, and functions of the human respiratory system. Our data were derived from video-recordings, semi-structured interview guides, and emotion diaries. Our data analysis showed that Susan constructed three epistemic emotions namely: frustration, anxiety, and joy while in the group. While frustration and anxiety were negative epistemic emotions that facilitated her non-participation and passive participation to their small group modeling activity, the positive epistemic emotion of joy facilitated her active participation. She constructed frustration in the Lesson 1 because of combined feeling of self-incompetence, thus, having the fear of being judged by her groupmates. However, this was changed to anxiety in Lesson 2 which was attributed to her lack of confidence to share her scientific knowledge. Nonetheless, she was able to passively participate when she started to feel the sense of belongingness as her groupmates emphatically provided her opportunities to participate. When the third lesson ended, she was reminded by the teacher to just continue interacting as she had great improvement from the first lesson. Thus, she was able to construct the epistemic emotion of joy from Lesson 5 that extended until Lesson 6. With joy, her cognitive and social engagements with the group were filled with fun which changed the group atmosphere.

Keywords: constructed epistemic emotions, epistemic emotions, scientific modeling

INTRODUCTION AND RATIONALE OF THE STUDY

Early constructionists claim that emotions are linked to the relationships between and among interacting individuals which can affect their participation in the social system (McCarthy, 1994; Wallbott & Scherer, 1986; Parkinson, Fischer, & Manstead, 2005). While Boiger and Mesquita (2012) mentioned that there are various contexts in which emotions are constructed, they specifically mentioned the social construction of emotions within three contexts namely, moment-to-moment interactions, developing and ongoing relationships, and sociocultural contexts. The prominence of small-group activities in science classrooms allows these three contexts to exist that trigger students' momentary construction of epistemic emotions. However, despite the growing number of studies on emotions and cognitions, research in this area especially in scientific modeling is scant.

We draw the importance of this study in that epistemic emotions have significant roles in the students' cognitive lives by influencing various aspects in the learning process such as their actions, behaviors, and motivations when they are engaged in knowledge construction (Arango-Muñoz & Michaelian, 2014). From the previous claims on epistemic emotion and learning, we hypothesize that students' participation in scientific modeling depends on the various epistemic emotions they construct that facilitate their performance in the epistemic practice. With gaps in the research on how epistemic emotions can be constructed in specific social situations, we investigated how an inherently passive student would eventually participate while constructing certain epistemic emotions. These epistemic emotions were explored in terms of how they influenced Susan's pattern of participation in scientific modeling by focusing on the following questions:

1. What epistemic emotions were constructed by Susan across the scientific modeling lessons?
2. How was her participation in the small group activities influenced by her constructed epistemic emotions?

METHODS

To investigate our research question, we explored a fifth-grade classroom taught by one of the authors in a public elementary school in one of the cities of Gyeonggi-do Province in South Korea during the 2018-2019 academic year. The teacher has been teaching in this school for 12 years and was using scientific modeling as one of the strategies in science class. Prior to the study, Susan's homeroom teacher conducted observations on her cognitive and social participation in classroom activities for 10 months not only in science classes but also with other subjects. In our initial analysis, Susan was mostly active in other subjects except in science. When we documented their group interactions across the six lessons on scientific modeling of the human respiratory system, we also noted that she constructed three different emotions which facilitated three different patterns of participation when they were accomplishing the required tasks of the activity. These served as the unique characteristics of Susan as our case student.

The scientific modeling of the human respiratory system consisted of six lessons where students were asked to work and interact in small groups. The cognitive tasks in the scientific modeling activity that were performed by students in this study were designed to understand what kinds of epistemic emotions they constructed during the moment-to-moment interactions while they develop and establish on-going relationships with each other. They were then expected to cognitively participate by articulating their claims, grounds, and rebuttals towards their group members' explanations and eventually construct a model for the structure, functions, and mechanisms of the human respiratory system.

Using video-recordings, semi-structured interview guides, emotion diaries, and field notes, we investigated what epistemic emotions Susan constructed while performing the cognitive tasks of the lesson with her groupmates. These were analyzed and interpreted on how they influenced her participation in the group social and cognitive tasks. Analysis was conducted through iterative viewing of the video-recordings paying attention to the facial expressions, gestures, tone of voices, and words she used (Jaber, 2014). We noted her repetitive behaviors she displayed as a product of her constructed epistemic emotions while in the small group. These were then verified using other data such as the interview and the emotion diaries.

RESULTS AND DISCUSSION

From across the six lessons of their scientific modeling activity, Susan constructed the epistemic emotion of *frustration* (Lesson 1), *anxiety* (Lesson 2 to 3), and *joy* (Lesson 5 to 6). These constructed epistemic emotions facilitated the different patterns of her participation during their scientific modeling class. It is important to note that we were not able to report results for Lesson 4 because their tasks were mostly illustrating and tracing the path of air on the different organs of the human respiratory system.

According to her emotion diaries, Susan constructed *frustration* as an epistemic emotion in Lesson 1 which facilitated her feeling of disinterest in interacting with the other members of her group. During the post-lesson interview, she expressed difficulty on the tasks which made her frustrated (Excerpt 1). Another factor which contributed to Susan's construction of *frustration* was her low perception towards her cognitive ability relative to the other members of her group. We were able to note this from the transcript of the post-lesson interview when she said that she is not as good as them in science.

Excerpt 1	
1	Teacher-researcher: *(After watching the video recording)* Susan, you rarely talked in Lesson 1, when Jenny asked your arguments, you just spoke out without confidence. Is it because you felt frustrated as you have written in your emotion diary?
2	Susan: Yes. I think it was hard on several tasks.
3	Teacher-researcher: Can you tell me what was difficult?
4	Susan: I think it is difficult for me to tell my arguments based on claims and evidence. I've never talked like this before. That was pressure!
5	Teacher-researcher: I think you are still able to tell your opinions to your students. Aren't you? Nobody can prevent you from telling your opinions.
6	Susan: Right.
7	Teacher-researcher: Isn't it just because it's hard to say?
8	Susan: You know, Jenny and Sylvia are good at science. But I am not. I don't think that my opinion will be accepted easily by them. Maybe they will just refute my opinion. So, this makes it hard for me and it's frustrating.

In assessing Susan's participation, we adapted the concept of productive participation (Engle & Conant, 2002) which considers the student's willingness and expressions of interest to be involved socially and cognitively in any epistemic activity. In scientific modeling like in this study, these two are intertwined with each other because the activities such as speaking, listening, constructing, and argumentation calls for the students' social and cognitive participation. Thus, in our analysis of Susan's level of participation, though she was able to respond, we still classified her responses as non-participation since she only participated when deliberately asked. Most of the time, she was not attentive, isolated herself in the discussion by messing around and playing with her pencils, and her responses did not contribute to the completion of their cognitive tasks (Excerpt 2).

Excerpt 2	
1	Jenny: First we need to explain which body organ has to be included into the human respiratory system. And why these body organs are needed for breathing? Oh, but it's a little awkward. Say that using claims and evidence (laugh).
2	Sylvia: I can't speak any more like this. Anyway...
3	Jenny: I think we need a nose, because we get air through our nose ... we need a nose.
4	Sylvia: Right. We can let air through our nose and nose filters out dirty things from the air.

5	Jenny: Mike, Susan. Why don't you guys tell us something?
6	Sylvia: I think we need trachea. Because air can move in and out through the trachea.
7	Jenny: Susan, why don't you do something else and focus on it?
8	Susan: *(Keeps looking at other places, and she is messing with pencils)*
9	Sylvia: Lungs are important, I think.
10	Jenny: I agree with you, too.
11	Sylvia: Me too, it delivers air into the vessels, especially oxygen (reading textbooks). Jenny: What are the other body organs that can be found in the human respiratory system?
12	Sylvia: Let's write it on paper once. We need to get it organized.
13	Jenny: Hey, Susan (aggressively), why do you think you need lungs?
14	Susan: Um ... that's what it's like *(No confidence; no social and cognitive contribution).*
15	Jenny: Tell me. Why do you think your lungs are needed?
16	Susan: Um ... to live? *(No confidence; no social ang cognitive contribution).*

In Lesson 2, Susan constructed the epistemic emotion of *anxiety*. This was triggered by her perception that her answers will likely be judged by the other members of the group. She also had low perception of her low academic competence. This was observed from her utterances when she said that she was not sure whether her answers were right or wrong (Excerpt 3).

Excerpt 3	
1	Teacher: You said you felt anxious in Lesson 2. Can you tell me what made you feel anxious?
2	Susan: I felt anxious because I was not sure if what I knew was right or wrong.
3	Teacher: What exactly does that mean?
4	Susan: I felt anxious when I told my opinion to Jenny and Sylvia; I was not sure whether what I was saying was scientifically correct or not.
5	Teacher: *(Shows the video-recordings)* But I noticed that you were able to talk to Jenny and Sylvia in Lesson 2. What has changed compared to Lesson 1?
6	Susan: Yes. I think it was less difficult for Jenny and Sylvia to show me that they would accept everything I would talk about.

Based on our analysis, Susan showed passive participation in their small group activity in Lesson 2. Though she was approached emphatically and was given enough opportunities to engage, she was still not confident in her responses (Excerpt 4). However, compared to the dynamics of their small group in Lesson 1, Susan was able to contribute to the completion of their cognitive tasks when she was emphatically given enough chance to speak. This confirms the work of Jaber, Southerland, and Dake (2018), group empathy enhances participation in small group scientific modeling.

Excerpt 4	
18	Sylvia: But maybe we can breathe out because our brain commands us to breathe for ourselves. If not, will we die? So, I think we need to draw a brain as part of the human respiratory system.
19	Jenny: Um ... Susan. I want you to tell me what you think. Talk comfortably.
20	Susan: *(Without confidence)* I don't think we need a brain in the drawing because they have no air pockets, lungs, instead. So, it's better not to draw.
21	Jenny: You may think so *(showing epistemic empathy).* But what if we don't breathe because we don't have a brain? So, I think we should draw the brain, also the body (to include lungs). What do you think? Susan?
22	Susan: *(With a little trembling voice)* So the body must be drawn. Because it's connected to the head, and to the respiratory organs.
23	Jenny: *(Smiling)* But didn't you say we don't need it? Anyway...
24	Sylvia: I think we can draw the head because we need the nose.

Based on Susan's emotion diaries, she constructed *joy* as an epistemic emotion from Lesson 5 to Lesson 6. We observed this with her smiles and playful behaviors when we scanned the video of their small group scientific modeling activity. In the interview, she said that the other members of their group continued to encourage and praise her. She specifically mentioned that her answers were not judged, and she was welcomed to do some playful acts while manipulating their model (Excerpt 5). She was then able to confirm that they do not have negative impressions on her which made her boost her confidence.

Excerpt 5	
1	Teacher-researcher: *(Shows the video-recordings)* Susan, at the beginning of Lesson 5, you seemed so comfortable enough to play with your groupmates. It was surprising to see that compared to the first period. Is there any reason why you acted like that?
2	Susan: I think it's because Jenny and Sylvia encouraged me and accepted whatever I said. It's a lot easier to tell my opinion to them. Actually, Jenny and Sylvia aren't uncomfortable with me, but at least in class, I think I should tell them something correct. You know, Jenny and Sylvia are good students. However, in Lesson 5, even if I joked, Jenny and Sylvia laughed with me without any reproach.

Aside from what was indicated in her emotion diaries, we also interpreted her groupmates' positive reactions such as their smiles and repetition of her actions as a scaffold to enable Susan to actively engage in their group activities (Excerpt 6). The emotion of joy that she epistemically constructed resulted to ease in dealing with others, increased her motivation in participating in the knowledge construction, and boosted her confidence in expressing her ideas. This supports the claims of Pekrun, Elliot, and Maier (2009) that positive emotional experience such as joy may improve learning outcomes; and in our study, the joy that Susan constructed led her to actively participate in the epistemic practice.

Excerpt 6	
29	Sylvia: *(Manipulating the syringe)* This action is air coming in, this action is air coming out, Susan, could you manipulate the syringe *(smiling)*?
30	Susan: *(Manipulates the syringe)* Ladies and gentlemen, listen up, I can make sound like this. Air comes in and out. Shuuuuuu…
31	Jenny and Sylvia: *(Smiling due to Susan's action and words)*
32	Susan: This is air coming in (Shuuuuu). Ladies and gentlemen, that's it! Thank you.
33	Jenny, Sylvia: *(Smiling due to Susan's action and words)*
34	Sylvia: *(Following Susan's action).* This is air coming in. Shuuuuuu…

In summary, two main factors influenced Susan's construction of epistemic emotions: 1) task familiarity, and 2) sense of acceptance. These were further affected my several underlying factors such as prior knowledge, argumentation skill, exposure to argumentation, motivation to do the tasks, curiosity about the lesson, difficulty level on the task, and certainty of ideas with varying influences in each of her engagements in the group activity in each lesson (Fig. 1).

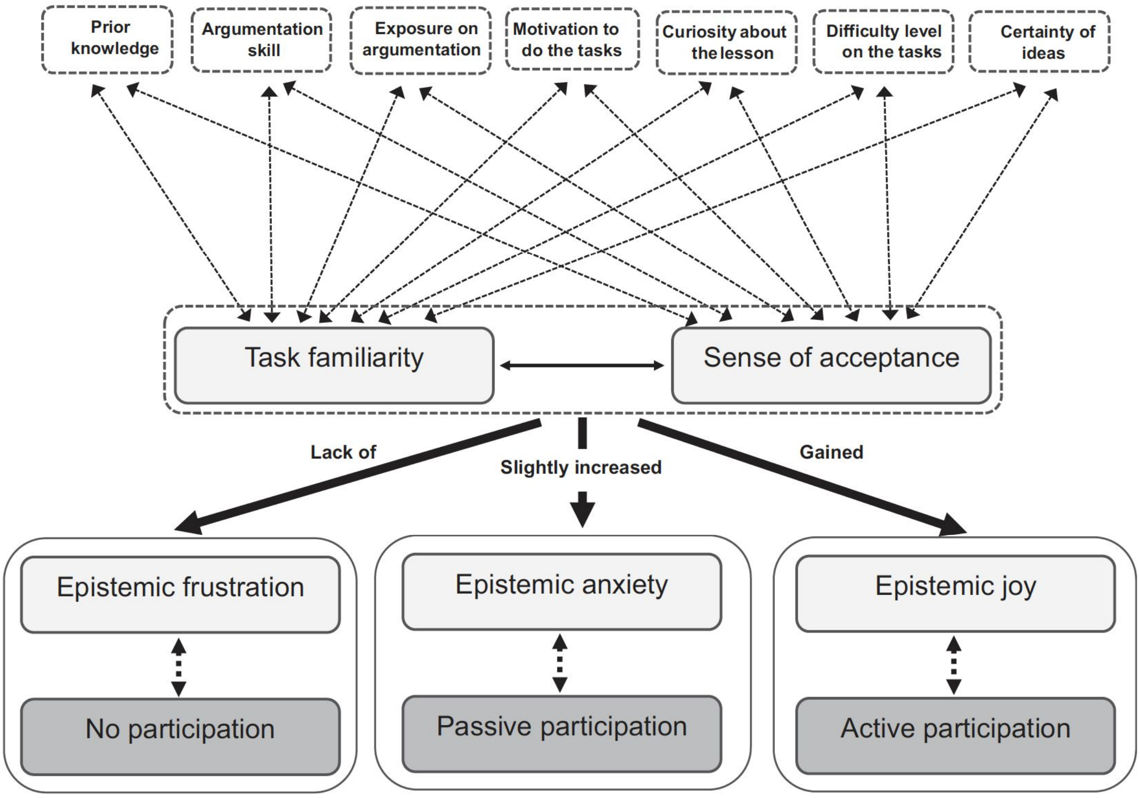


Figure 1. Reciprocal relations of the factors that contributed to Susan's construction of epistemic emotions and patterns of participation in the scientific argumentation and modeling.

In Lesson 1, Susan constructed epistemic *frustration* because of her unfamiliarity with the epistemic practices of science argumentation which was their main task in their group activity. In Lessons 2 and 3, she was able to slightly recover from unfamiliarity which relieved her discomfort in interacting with her groupmates. However, her lack of prior knowledge about the lesson topic and her lack of argumentation skill were not enough to motivate her participation; thus, facilitating her construction of epistemic *anxiety*. In Lessons 5 and 6, Susan's increased exposure boosted her confidence which resulted to her construction of epistemic joy that eventually facilitated her active participation. This is in combination with her improved sense of acceptance from her groupmates.

Previous studies which used the cognitive-motivational model distinguished emotions according to their performance effects: positive-activating, positive-deactivating, negative-activating, and negative-deactivating (Pekrun et al., 2002; Pekrun, Elliot, & Maier, 2006). In this study, we noted that epistemic *joy* was a positive-activating emotion while epistemic *frustration* and epistemic *anxiety* were negative-deactivating emotions. According to Pekrun et al. (2002), positive-activating emotions such as epistemic *joy* reinforce motivation for learning. This was observed when Susan displayed task-related enjoyment. We particularly observed this when she had fun in explaining and summarizing their group arguments and in her relational processing of information when she was actively engaged with Jenny and Sylvia in explaining and manipulating the syringe as a scientific model. We therefore confirm previous studies which hypothesized that positive epistemic emotions drive students' collaboration in group activities (Nicolaou, Evagorou, & Lymouridou, 2015). Her epistemic *joy*, which was triggered by her familiarity on their tasks and her perceived sense of acceptance by Jenny and Sylvia helped her to pay attention to the discussion and eventually engage with unsolicited arguments. On the contrary, her epistemic *frustration* and epistemic *anxiety* limited her participation in scientific argumentation and activated her avoidance due to her fear of committing mistakes. Her desire to engage in the small group argumentative discussions was affected by her perception that her ideas will be rejected resulting from her lack of prior knowledge. This sense of rejection led her to avoid the task even though she was deliberately asked by her groupmates to tell her opinions.

To assess the impact of her negative epistemic emotions in relation to the concept of productive participation (Engle & Conant, 2002), these negative epistemic emotions affected not only herself but also her small group cognitive and social performance. Instead of voluntarily sharing her ideas which may increase their opportunities for scientific argumentation in their small group, her limited participation due to her negative emotions suppressed their learning opportunities in the group. However, more than the group members who were affected, we claim that her emotions had more negative impacts on her because even without her participation, the other members of her group were able to proceed and accomplish their tasks. This confirms previous studies report that an individual's emotions still have great impacts compared to group emotions (Boekaerts, 2007). Thus, socio-cognitive interactions, though may seem to trigger construction of collective or group epistemic emotions, each member's construction of emotions appears to be more influential in facilitating their patterns of participation.

CONCLUSIONS AND IMPLICATIONS

The results of the study addressed the multi-faceted effects of Susan's prior knowledge and skills, level of understanding of the science concepts, how she perceived herself relative to her classmates, and the ongoing and developing into her construction of epistemic emotions. These results indicated that epistemic emotions and the socio-cognitive interactions in a small group have reciprocal relations to a student's pattern of participation in a small group modeling activity. Our analyses contribute to the ideas of the constructionists' views of emotions which emphasized that instead of being prewired from birth, emotions are derived from the changing dynamics of interaction and relational patterns in a social environment where an individual belongs (Boiger & Mesquita, 2012). Thus, we report the possibility that the momentary nature of epistemic emotions may facilitate different patterns of participation in the classroom activities.

In this study, Susan, represented the case of a passive student in science, who showed no participation and passive participation in the early stages of their lessons when she constructed negative epistemic emotions. However, when she constructed the positive epistemic emotion of *joy* as soon as she was able to resolve her difficulties, she was able to show active participation especially when she received support from her groupmates.

The primary implication of this study is provision of evidence on how epistemic emotions are socially constructed. Our analysis emphasized that the construction of epistemic emotions is momentary and unconscious which may shift the nature of students' participation during social interaction. This supports earlier studies that contingent on the different activities that occur in the social system is the emergence and construction of epistemic emotions (Boiger & Mesquita, 2012). It is therefore necessary for teachers to make efforts in identifying these emotions at least through students' emotion diaries.

Thus, we recommend that for teachers to harness the significance of small group modeling, they must be aware of the nature of the tasks and the students' prior knowledge which may influence their intrinsic motivation and excitement. Moreover, they should also be mindful and be ready to intervene especially when they have students who are inherently active with the tendency to dominate the small group activity. In our study, the teacher allotted some time to conduct post-session processing after their group activities to remind the students give students in the group that each one has to be given opportunities to contribute while displaying emotional empathy on each other's opinion. We therefore suggest that teachers should be mindful of these scenarios by assigning student roles to ensure the participatory construction of knowledge.

With the contention that epistemic emotions are implicated during cognition and learning, we also recommend that future research explore the possibility of shared constructed epistemic emotions. Studies can focus on how and what factors contribute to the shared constructed epistemic emotions in a small group of students as they engage in an epistemic inquiry. The differences in the socially constructed epistemic emotions which was also observed in this study can be given attention by exploring other triggering factors for each member of the small group.

REFERENCES

Arango-Muñoz, S., & Michaelian, K. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries, 2*, 97–122.

Boekaerts, M. (2007). Understanding students' affective processes in the classroom. In P. A. Schutz, & R. Pekrun (Eds.). *Emotion in Education* (pp. 37–56). https://doi.org/10.1016/B978-012372545-5/50004-6

Boiger, M., & Mesquita, B. (2012). Construction of emotion in interactions, relationships, and cultures. *Emotion Review, 4*(3), 221-229.

Engle, R. A. & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners' classroom. *Cognition and Instruction, 20,* 399-483.

Jaber, L. (2014). *Affective dynamics of students' disciplinary engagement in science*. Doctoral dissertation, Tufts University, MA.

Jaber, L., Southerland, S., & Dake, F. (2018). Cultivating epistemic empathy in preservice teacher education. *Teacher and Teacher Education, 72,* 13-23.

McCarthy, E. D. (1994). The social construction of emotions: New directions from culture theory. Sociology Faculty Publications. 4. https://fordham.bepress.com/soc_facultypubs/4

Nicolaou, C., Evagorou, T. M., & Lymbouridou, C. (2015). Elementary school students' emotions when exploring an authentic socio-scientific issue through the use of models. *Science Education International, 26*(2), 240–259.

Parkinson, B., Fischer, A., & Manstead, A. S. R. (2005). *Emotion in social relations: Cultural, group, and interpersonal processes.* Philadelphia, PA: Psychology Press.

Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology,* *101*, 115–135.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of quantitative and qualitative research. *Educational Psychologist, 37*, 91–106. https://doi.org/10.1207/S15326985EP3702_4

Wallbott, H. G., & Scherer, K. R. (1986). How universal and specific is emotional experience? Evidence from 27 countries on five continents. *Social Science Information, 25*(4), 63-95. https://doi.org/10.1177/053901886025004001

PRELIMINARY RESULTS OF A PARAMETRIC ANALYSIS OF EMOTIONS IN A LEARNING PROCESS IN SCIENCE

Federico Agen[1], Ivan Ezquerra-Romano[2] and Angel Ezquerra[1]

[1]Science, Social Science and Mathematics Education, Universidad Complutense de Madrid, Spain.

[2]Institute of Cognitive Neuroscience, University College London (UCL), London, UK.

In the last decades, several studies have highlighted the importance of emotions in the teaching and learning process. The classroom is considered an emotional place, where the learning is influenced by cognitive and emotional-motivational mechanisms. Classically, emotions have been classified in discrete categories. Furthermore, in educational settings, it is possible to evaluate dimensional categories as engagement and attention. According to this vision, we designed an activity to analyse emotions and their flow when students are involved in an inquiry-based activity. To avoid limitations of self-reports and observational methods, we evaluated emotions with an automatic facial coding system. This system detects facial human expressions using facial reference points and classifies their emotional value parametrically. The data shows different flows for each affective parameter. Thus, we obtained a constant high level of attention and intense engagement along the whole activity. Moreover, joy and surprise flows showed a global presence higher than negative emotions. Four parameters' flow graphics are related to characteristic educational behaviour. This work opens to the possibility of objective parametric evaluations of the emotional component in the teaching-learning process.

Keywords: Emotion, Science Education, Conceptual change.

INTRODUCTION AND THEORETICAL FRAMEWORK

In Education emotions may be described as appraisals and reactions to the information received from the context, whose intensity depends on subjective evaluations, influenced by personal prior knowledge beliefs and priorities (Dávila et al., 2021; Graesser, 2020; Harley et al., 2017; Rubin & Talarico; 2009). Emotions may occur due to evocation of events that happened in the past or by anticipating possible future situations (Damasio & Carvalho, 2013; Hutchinson & Barrett, 2019; Pekrun et al., 2014; Russell, 2003). Thus, antecedents as academic successes and failures, may shape students' emotions. Conversely, emotions may impact outcomes, determining a reciprocal influence (Murphy, 2019; Pekrun et al., 2017).

Emotions influence in different ways experiences, strategies, and attitude of students toward learning (Borrachero et al., 2014; Murphy et al., 2019). They may be positive (happiness, hope, etc..) or negative (anxiety, frustration, etc..). In general, positive emotions (such as enjoyment, hope, satisfaction, self-confidence) support the teaching-learning process positively (Pekrun et al., 2017). On the contrary, negative emotions (such as boredom, confusion, frustration, and hopelessness) tend to have negative influences on learning and are negatively related to achievement (Murphy et al., 2019; Marcos-Merino, 2019; Pekrun et al., 2014).

Moreover, the emotional state can be described dimensionally. Dimensions describes emotional experiences along a continual variation of parameters as pleasure and arousal, directly related to core affect (Deckert et al., 2019; Plass et al, 2019). Anyway, dimensions may represent a superior level that includes emotion discrete categories (Loderer et al. 2019). Some prominent dimensions in educational context are attention and engagement (Graesser, 2020; Harley et al., 2017; Loderer et al., 2019; Mrkva et al., 2019).

Introduction to parametric study of emotions

Basically, up to the present days, scientific research in Education has been focused on observational and self-reports methods (Azari et al., 2020; Harley et al., 2019; Loderer et al., 2019; Meindl et al., 2018; Pekrun, 2006). These methodologies, for assessing personal qualities, are the most common approaches in research (Duckworth & Yeager, 2015). They have the capacity to get outcomes in a cheap, quick and versatile way of emotion recognition (Engelman & Bannert, 2019).

Anyway, self-reports are difficult to edit and the declarations of the participants may be affected by biases, as consequences of incapacity to self-estimate or express correctly their own emotions (Engelman & Bannert, 2019; Goetz et al., 2016; Izard, 2009; Pekrun, 2006). As far as observer-reports, coder experts need a long and intense practice to achieve reliable data by observation (Barrett et al., 2019).

In the last decades, the continuous development of technology offers different methodologies and analytical instruments to identify the emotional process. These systems collect and analyse data related to the brain and nervous system. Some among the most investigated are EEG, heart rate, skin conductance, fMRI and eye tracking. Furthermore, recent advanced included no-intrusive techniques to automate facial expression recognition systems using cameras or webcams (Darvishi et al., 2021; Monkaresi et al., 2017), now commercially available for scientific research with high level of reliability (Stöckli et al., 2018, Küntzler et al., 2021). This is relevant in education, considering that the movements of the facial muscles almost always accompany an emotional state, which can be related to discrete emotions and affective dimensions. Thus, the analysis of facial expressions is one of the most appropriate automatized techniques to estimate emotions and behaviours in class (Calado et al., 2017).

Facial automatic detection systems can achieve an appropriate and accurate postural, head movement and facial expression coding recognition around 90%, depending on the conditions of clear and correct illumination of the participant's face (Benitez-Quiroz et al. 2017). That gives us the chance to bring to bear a dynamic perspective on emotional changes over a period of time. Thus, it is possible to analyse profiles of emotional response and behaviour occurring in a given situation (Gross, 2015; Kuppens et al., 2009; Kuppens & Verduyn, 2015).

The first step of the automatic facial expression recognition was to detect the face of the subjects (Kulkarni et al., 2021). For that, the software iMotions (2018) utilised the algorithm Viola Jones Cascaded Classifier (Viola & Jones, 2004). Successively, to estimate the facial expression recorded, it used an automatic coding system based on Facial Action Coding System (FACS) (Barrett et al., 2019; Ekman & Friesen, 1978; Keltner et al., 2019). Eventually, the Affectiva AFFDEX algorithm SDK 4 (Affectiva, 2015, Boston, MA) correlated the facial expressions to the affective states. This system can detect head orientation (yaw, pitch, roll); interocular distance; 34 facial landmarks; 14 facial expression metrics.

In detail, to describe facial movement, the software uses algorithms to detect landmarks as brows, mouth corners, etc., as well as groups of landmarks. When these reference points change their relative position, due to a change of respondent's expression, the system evaluates the new facial configuration in terms of affective metrics. Each movement corresponds to an Action Unit (AU). One or more AUs describe an emotional facial expression based on FACS. It allows to assess basic emotions: Anger, Sadness, Joy, Surprise, Disgust, Contempt and Fear; moreover, affective dimensions as Attention, Valence and Engagement. When an emotional event occurs, it generates an emotional episode that is evident in the change of facial configuration, ending when it goes back to its baseline level (Kuppens et al., 2015). Because respondents differ in their natural expression, the Affectiva algorithm applies a rolling baseline on the neutral expression of the respondent. This process keeps into account the frames preceding and following the current frame, and calculates changes. Each frame get an assigned score, depending on facial expression recognition and its intensity, from the absence of expression (0%) to an expression fully present (100%).

In this work, based on an artificial intelligence system for facial expression recognition, we have set out to develop and apply an experimental design which would allow us to collect and study the emotional and behavioural dynamics in a science education activity.

RESEARCH METHOD AND DESIGN

With the aim to investigate emotional and behavioural dynamics in science education, we proposed an inquiry activity to 24 teaching students (15 women and 9 men) attending the Master in Secondary Education, at School of Education, Complutense University of Madrid. The participants had to predict the contents of a box, with dimensions of 9x6x20 cm. It contained some euro coins: two of 1 cent; one of 5 cents; one of 10 cents; one of 20 cents; one of 1 euro. All these coins moved freely inside it. The participants could not open the box or break it. That is, they had to make use only of their scientific-technical knowledge, such as observing, testing hypotheses, drawing conclusions, etc. They could use some magnets. A similar activity was proposed among others by Lederman & Abd-El-Khalick (1998) and Haber-Schaim et al. (1979).

The participants were divided into pairs. Only one student of each pair had to guess the contents, and they were video-recorded. The other students observed their peers' activity. They checked the right operation of the camcorder, and they warned their peers to remain inside the video framing. In this way they learnt the operational best practice. Each HD video camera was placed on a tripod in front of each observed student, at a distance of one metre to obtain the best recording view of the face and upper body. The activity lasted twenty minutes. We divided the session into ten periods of two minutes each. After each period, students filled a form in which they reported their emotions. Nevertheless, here, we only describe the emotional dynamics obtained from the video recordings.

After data collection, we devised the protocol to process and analyse the data. The videos recorded during the activity were saved and named with a specific code. Later, they had been edited to prepare them for the analysis. Specifically, they were synchronized with the start, the frames recorded outside of the activity's duration were cut off, as well the pause intervals (time utilized by the students to take notes), then the remaining parts were merged. The videos were imported and processed by iMotions® program.

The total processed data constituted our initial signal for analysis, it consisted of 60.856.164 entries, determined by 141 entries per frame. The frames analysed were 35.967 for each respondent (12).

RESULTS

Our preliminary results indicated different dynamics for each affective parameter. Throughout the ten time periods (Table 1), we observed Attention with the highest global average presence value (69,4%), then Engagement (20,8%), Joy (4,8%) and Surprise (3,2%). Other emotions were quite lower (< 1%). Basically, four emotional states were prevalent for all participants. The Standard Deviation (SD) was restrained to a range from 0,05% (Sadness) to 3,91% (Attention). It seems to indicate that emotional states, experienced by the students, maintained their average presence percent along the whole activity. Effectively, the analysis of the parameters in particular with high presence, Attention and Engagement, indicated that all the students were engaged with high attention carrying out the activity.

Table 1. Percent of time for each of the 9 selected parameters, they were measured by the system for each of the periods. "Mean" indicates the mean value of each emotion for all periods within the 12 respondents.

	Anger	Sadness	Disgust	Joy	Surprise	Fear	Contempt	Engagement	Attention
Period 1	0,00%	0,10%	0,30%	9,40%	2,90%	0,30%	1,60%	25,10%	60,30%
Period 2	0,90%	0,00%	0,30%	7,50%	1,50%	0,50%	0,90%	22,30%	66,10%
Period 3	0,70%	0,00%	0,10%	3,90%	1,90%	0,10%	0,90%	17,20%	75,10%
Period 4	0,00%	0,00%	0,20%	4,60%	2,20%	0,40%	1,00%	20,60%	71,00%
Period 5	0,10%	0,00%	0,30%	3,60%	3,10%	1,10%	1,30%	19,40%	71,30%
Period 6	0,10%	0,00%	0,10%	5,00%	4,50%	0,60%	0,40%	21,00%	68,30%
Period 7	0,20%	0,10%	0,30%	2,20%	3,90%	0,50%	0,70%	19,40%	72,70%
Period 8	0,00%	0,00%	0,30%	3,90%	4,20%	0,50%	0,70%	21,90%	68,50%
Period 9	0,10%	0,00%	0,20%	3,30%	3,40%	0,80%	0,50%	19,60%	72,00%
Period 10	0,00%	0,10%	0,10%	4,50%	4,30%	0,50%	0,40%	21,70%	68,30%
Mean	0,20%	0,00%	0,20%	4,80%	3,20%	0,50%	0,80%	20,80%	69,40%
SD	0,30%	0,05%	0,09%	2,02%	1,01%	0,26%	0,37%	2,03%	3,91%

We analysed the graphics of the parameters, focusing specifically on the dynamics of the prevalent four. Then, we observed whether there were profiles which were shared by the respondents. The analysis indicated three different profiles in parameters' graphics. In Figures 1 (Surprise and Engagement) and Figure 2 (Attention) we show the graphics of Respondent 10 as an example.

Profile 1 was correlated to the first part of the activity (in this case till about 6 minutes) and presented Surprise (orange line in Figure 1) with very low presence or absence of peaks. Engagement (light blue line in Figure 1) was characterised by high variability, with sharp peaks and low interval average presence (15%). Attention (Figure 2) was very high for the initial first minute, but then (1-6 minutes interval) it showed high variability with repeated rise and fall, for an interval average presence of 76%.

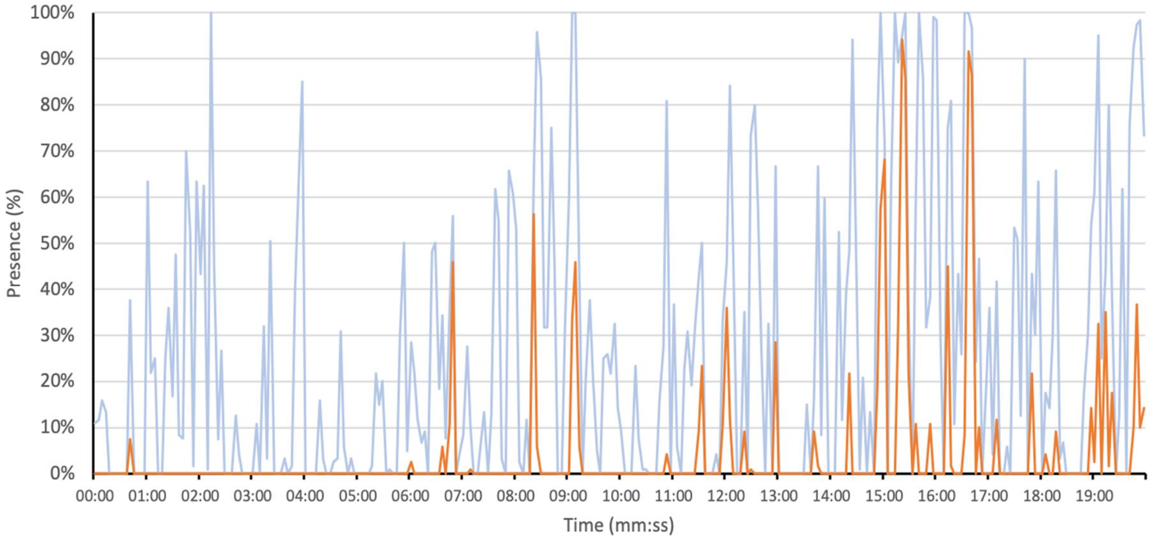


Figure 1. Surprise (orange) and Engagement (light blue) (Respondent 10).

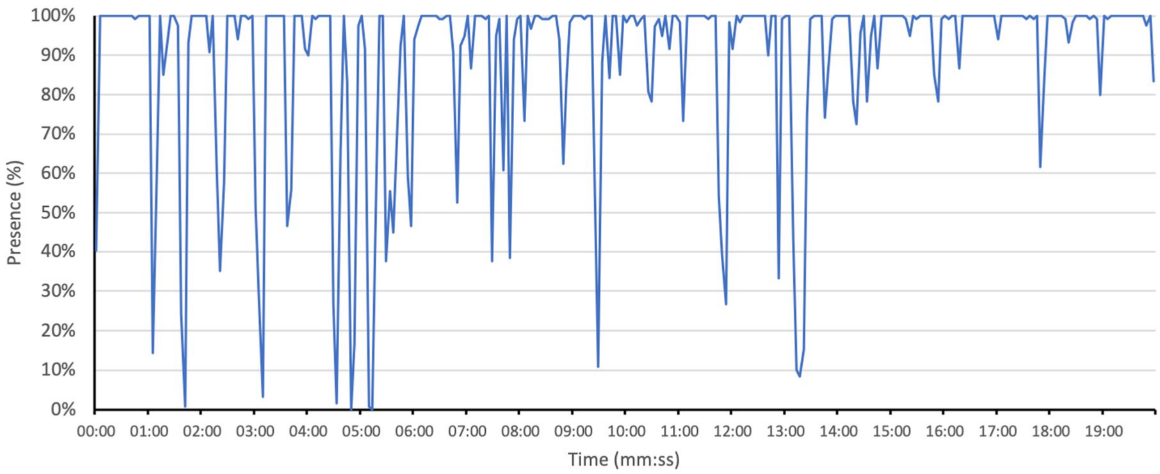


Figure 2. Attention (Blue) (Respondent 10).

Profile 2 was correlated with the central part of the activity (in this case about 6-14 minutes interval). It was characterized by a sequence of isolated sharp peaks of Surprise, of medium and low intensity. It showed covariation with Engagement, that presented variability of peaks succeeding with different intensity and frequency, with an average presence (24%), higher than the previous interval. Attention (Figure 2) presented high presence with variability less evident than the previous profile and showed peaks with either steep or smooth slopes. It showed short tracts with inertia and an interval average presence (89%) higher than the preceding interval.

Profile 3 referred to the last part of the activity (in this case about 14-20 minutes interval). Surprise (Figure 1) showed isolated peaks of different intensity, clustered in the central part of the interval, with some peaks of high intensity of presence (70-90%). It still presented covariation with Engagement, that increased the interval average presence (42%) with higher sharp peaks frequency and roller-coaster trend. Attention (Figure 2) showed an inertia maintained till to describe a plateau, where Attention persisted for several seconds or few minutes at 100% presence, thus achieving the highest interval average presence (97%).

We evaluated the presence and the sequence of the profiles and we could observe different duration depending on the participants. Thus, we could propose common patterns within the respondents, depending on the profiles they showed. This introductory analysis indicated relations of the parameters' patterns (achieved from data analysis) with students' educational actions (inferred from observing the recorded videos).

Pattern 1 showed the sequence of the profiles 1, 2 and 3. Basically, Profile 1 was present at the beginning of the activity, Profile 2 characterized the central part of the time-line, whereas Profile 3 was prevalent on the last part of the activity. This pattern was shared among Respondents 1, 2, 9, 10, 11. This preliminary study of the students' educational actions showed their high capacity of concentration, problem solving and to apply systematically the hypotheses elaborated.

Pattern 2 showed a short tract with the Profile 1 and alternation of Profile 2 and 3, with short (2-3-2-3) or large (2-3-2-3-2-3) sequences, with different intervals for each participant (3, 4, 5, 7, 8). They also showed interest and implication, but more doubts and uncertainty than the previous group.

Pattern 3 didn't present any of the mentioned profiles (Respondents 6, 12), due to reduced engagement and the lack of characteristic trend of Attention and Surprise described for the profiles above. They showed difficulty to elaborate and apply strategy to resolve the task.

This preliminary study seems to point at the existence of emotional dynamics linked with the students' educational behaviours. Nevertheless, more time and data processing are necessary to establish more robust relationships.

DISCUSSION AND CONCLUSIONS

This study indicates the possibility to evaluate parametrically the emotions during an educational activity, overcoming some self-report or observational limitations. We had the possibility to continuously follow the dynamic of students' affective dimensions and emotions by using a facial recognition system.

The data showed the parameters of Attention and Engagement predominant throughout the activity. Positive emotions, Joy and Surprise, displayed global average presence higher than negative emotions. It indicated that the activity was carried out with motivation, implication and positive attitude by the students. Furthermore, the restrained standard deviation implied that the general trend of the different emotional states was consistent throughout the activity.

The analysis of Attention, Engagement and Surprise dynamics permitted to elaborate three graphical profiles. The distinct presence and sequence of these profiles, along the task, drew three patterns shared by different groups of participants. It is remarkable that each pattern's group of students showed a characteristic educational behaviour, related to observation, reflection, systematic exploration and application of strategies. For the students it implied a different ability to carry out the task, such as to develop and test hypotheses, to avoid distraction, capacity of concentration and perseverance.

Basically, the profiles and patterns found are peculiar to this activity. Different tasks would imply rather distinct profiles and patterns classification. Anyway, this study confirms the correlation between emotions and educational behaviours, and encourages us to expand the research to other educational activities.

Concerning the teaching-learning process, this preliminary study underlines the importance of teachers' capacity to correlate educational actions with the affective states that students go through, when they are involved in an educational activity. Evidently, the future teachers cannot still ignore the emotional dynamics in class. Thus, it should be an important part of their training.

It is worth considering some limitations, such as the difficulty for some respondents to avoid covering the face or turn it on a side over a suitable angle, aligned with the camera, to be correctly detected by the system, for the whole time of the activity. Moreover, the large quantity of data implicated the necessity of managing a notable noise reduction. Anyway, this work opens up the possibility for objective parametric evaluations of emotional components during the teaching-learning process.

ACKNOWLEDGEMENTS

This work has been funded by Project: "Identification of scientific contexts in society. Tools for teachers and citizens (SCIxSOC)" (RTI2018-094303-A-I00) of the Ministry of Science, Innovation, and Universities (Spain) in the framework of the National R&D&I Program Oriented to the Challenges of Society, State Plan for Scientific and Technical Research and Innovation 2019-21.

REFERENCES

Azari, B., Westlin, C., Satpute, A.B., Hutchinson, J.B., Kragel, P.A., Hoemann, K., Khan, Z., Wormwood, J.B., Quigley, K.S., Erdoğmuş, D., Dy, J.G., Brooks, D.H., & Barrett, L.F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports, 10*. https://doi.org/10.31234/osf.io/egh2t

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Benitez-Quiroz, C. F., Srinivasan, R., Feng, Q., Wang, Y., & Martinez, A. M. (2017). EmotioNet Challenge: Recognition of facial expressions of emotion in the wild. *ArXiv, abs/1703.01210*.

Borrachero, A. B., Brígido, M., Mellado, L., Costillo, E., & Mellado, V. (2014). Emotions in prospective secondary teachers when teaching science content, distinguishing by gender. *Research in Science & Technological Education,* *32*(2), 182–215. https://doi.org/10.1080/02635143.2014.909800

Calado, J., Luís-Ferreira, F., Sarraipa, J., & Jardim-Gonçalves, R. (2017). A Framework to Bridge Teachers, Student's Affective State, and Improve Academic Performance. *Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition. Volume 2: Advanced Manufacturing.* Tampa, Florida, USA. November 3–9, 2017. V002T02A042. ASME. https://doi.org/10.1115/IMECE2017-72000

Damasio, A.R., & Carvalho, G.B. (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience, 14*, 143-152.

Darvishi, A., Khosravi, H., Sadiq, S., & Weber, B. (2021). Neurophysiological measurements in higher education: a systematic literature review. *International Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00256-0

Dávila, M. A., Cañada, F., Sánchez-Martín, J., Airado, D., & Mellado, V. (2021). Emotional performance on physics and chemistry learning: the case of Spanish K-9 and K-10 students. *International Journal of Science Education, 43*(6), 823–846. https://doi.org/10.1080/09500693.2021.1889069

Deckert, M., Schmoeger, M., Auff, E., & Willinger, U. (2019). Subjective emotional arousal: an explorative study on the role of gender, age, intensity, emotion regulation difficulties, depression and anxiety symptoms, and meta-emotion. *Psychological Research, 84*, 1857 - 1876.

Duckworth, A.L., & Yeager, D.S. (2015). Measurement Matters. *Educational Researcher, 44*, 237 - 251.

Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.

Engelmann, K., & Bannert, M. (2019). Analyzing temporal data for understanding the learning process induced by metacognitive prompts. *Learning and Instruction, 101205*.

Goetz, T., Sticca, F., Pekrun, R., Murayama, K., & Elliot, A.J. (2016). Intraindividual relations between achievement goals and discrete achievement emotions: An experience sampling approach. *Learning and Instruction, 41*, 115-125.

Graesser, A. C. (2020). Emotions are the experiential glue of learning environments in the 21st century. *Learning and Instruction, 70*, 101212. https://doi.org/10.1016/j.learninstruc.2019.05.009

Gross, J.J. (2015). The Extended Process Model of Emotion Regulation: Elaborations, Applications, and Future Directions. *Psychological Inquiry, 26*, 130 - 137.

Haber-Schaim, U. y otros. (1979). *Curso de introducción a las ciencias físicas*. Barcelona: Reverté.

Harley, J. M., Jarrell, A., & Lajoie, S. P. (2019). Emotion regulation tendencies, achievement emotions, and physiological arousal in a medical diagnostic reasoning simulation. Instructional *Science, 47*, 151–180. https://doi.org/10.1007/s11251-018-09480-z

Harley, J.M., Lajoie, S.P., Frasson, C., & Hall, N. (2017). Developing Emotion-Aware, Advanced Learning Technologies: A Taxonomy of Approaches and Features. *International Journal of Artificial Intelligence in Education, 27*, 268-297. https://doi.org/10.1007/s40593-016-0126-8

Hutchinson, J.B., & Barrett, L.F. (2019). The Power of Predictions: An Emerging Paradigm for Psychological Research. *Current Directions in Psychological Science, 28*(3), 280–291. https://doi.org/10.1177/0963721419831992

iMotions Biometric Research Platform. (2018). [computer software], (8.1 Version). iMotions A/S, Copenhagen, Denmark.

Izard C. E. (2009). Emotion theory and research: highlights, unanswered questions, and emerging issues. *Annual review of psychology*, *60*, 1–25. https://doi.org/10.1146/annurev.psych.60.110707.163539

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression: Advances in Basic Emotion Theory. *Journal of nonverbal behavior, 43*(2), 133–160. https://doi.org/10.1007/s10919-019-00293-3

Kulkarni, K., Corneanu, C.A., Ofodile, I., Escalera, S., Baró, X., Hyniewska, S.J., Allik, J., & Anbarjafari, G. (2021). Automatic Recognition of Facial Displays of Unfelt Emotions. *IEEE Transactions on Affective Computing, 12*, 377-390.

Küntzler, T., Höfling, T.T., & Alpers, G.W. (2021). Automatic Facial Expression Recognition in Standardized and Non-standardized Emotional Expressions. *Frontiers in Psychology, 12*.

Kuppens, P., & Verduyn, P. (2015). Looking at Emotion Regulation Through the Window of Emotion Dynamics. *Psychological Inquiry, 26*, 72 - 79.

Kuppens, P., Stouten, J., & Mesquita, B. (2009). Individual differences in emotion components and dynamics: Introduction to the Special Issue. *Cognition and Emotion, 23*, 1249 - 1258.

Lederman, N., & Abd-El-Khalick, F. (1998). Avoiding de-natured science: Activities that promote understandings of the Nature of Science. In W. F. McComas (Ed.), *The Nature of Science in science education* (pp. 83–126). Dordrecht: Springer.

Loderer, K., Pekrun, R., & Jan L. Plass J.L. (2019). Affective Foundations of Game-Based Learning. In Plass, J.L., Richard E. Mayer, R.E., & Homer, B.D. (Ed.), *The Handbook of Game-based Learning*. MIT Press.

Marcos-Merino, J. M. (2019). Análisis de las relaciones emociones-aprendizaje de maestros en formación inicial con una práctica activa de Biología. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias, 16*(1), 1603. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i1.1603

Meindl, P., Yu, A., Galla, B.M., Quirk, A., Haeck, C., Goyer, J.P., Lejuez, C.W., D'Mello, S.K., & Duckworth, A.L. (2018). A brief behavioral measure of frustration tolerance predicts academic achievement immediately and two years later. *Emotion*.

Monkaresi, H., Bosch, N., Calvo, R.A., & D'Mello, S.K. (2017). Automated Detection of Engagement Using Video-Based Estimation of Facial Expressions and Heart Rate. *IEEE Transactions on Affective Computing, 8*, 15-28.

Mrkva, K., Westfall, J., & Van Boven, L. (2019). Attention Drives Emotion: Voluntary Visual Attention Increases Perceived Emotional Intensity. *Psychological Science, 30*, 942 - 954.

Murphy, S., MacDonald, A., Wang, C.A., & Danaia, L. (2019). Towards an Understanding of STEM Engagement: a Review of the Literature on Motivation and Academic Emotions. *Canadian Journal of Science, Mathematics and Technology Education, 19*(3), 304-320. https://doi.org/10.1007/s42330-019-00054-w

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R., Cusack, A., Murayama, K., Elliot, A.J., & Thomas, K. (2014). The power of anticipated feedback: Effects on students' achievement goals and achievement emotions. *Learning and Instruction, 29*, 115-124.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child development, 88*(5), 1653-1670.

Plass, J.L., Homer, B.D., Macnamara, A., Ober, T.M., Rose, M.C., Pawar, S., Hovey, C.M., & Olsen, A. (2019). Emotional design for digital games for learning: The effect of expression, color, shape, and dimensionality on the affective quality of game characters. *Learning and Instruction, 101194*.

Rubin, D. C., & Talarico, J. M. (2009). A comparison of dimensional models of emotion: Evidence from emotions, prototypical events, autobiographical memories, and words. *Memory, 17*(8), 802–808. https://doi.org/10.1080/09658210903130764

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. https://doi.org/10.1037/0033-295x.110.1.145

Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A.C. (2018). Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior Research Methods, 50*, 1446-1460.

Viola P. & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

ESERA 2021
Fostering scientific citizenship in an uncertain world
30 Aug - 3 Sep 2021
Organised by University of Minho, Braga, Portugal

STUDENTS' PERCEIVED COMPETENCE AS PREDICTOR OF THEIR FLOW EXPERIENCE DURING EXPERIMENTATION

Tim Kirchhoff, Matthias Wilde and Nadine Großmann

Bielefeld University, Bielefeld, Germany

Several studies have shown a decrease in student motivation in science education. To design measures to motivate students such as their flow experience variables that have an impact on flow experience must first need to be identified. Previous studies reveal that students' perception of competence (e.g., during experimentation) is crucial for their flow experience in science classes. According to the scientific discovery as dual search model, perceived competence during experimentation refers to the perception of competence during the three phases 'Search Hypotheses Space', 'Test Hypothesis', and 'Evaluate Evidence'. In this study, we investigated whether students' perceived competence during the aforementioned three phases of experimentation predicts their flow experience. To investigate these effects, 212 German students were taught in biology classes (age: M = 16.30 years, SD = 0.96 years; 59% female). The students performed two experiments. Immediately thereafter, students' perceived competence and their flow experience were assessed. Regression analysis revealed that students' perceived competence during the phases 'Test Hypothesis' and 'Evaluate Evidence' but not during the phase 'Search Hypotheses Space' predicted their flow experience. The more competent the students felt, the higher their flow experience. With these results, specific support measures for experimentation processes can be designed.

Keywords: scientific experimentation, perceived competence, flow experience

RATIONALE

Research demonstrates that students in Germany are losing their interest and motivation in science education (e.g., Krapp & Prenzel, 2011; Organisation for Economic Co-operation and Development, 2016), leading to many studies focusing on how to counteract this decrease. Science education employs several hands-on activities such as experimentation that might have positive effects on students' motivation (Hofstein & Lunetta, 2004). However, the process of experimentation can be rather complex. A wide range of skills and knowledge is required (Bruckermann, Arnold, Kremer, & Schlüter, 2017; Klahr & Dunbar, 1988), which students often lack (Baur, 2018; Chinn & Malhorta, 2002; Etkina, Murthy, & Zou, 2006; Hammann, Phan, Ehmer, & Bayrhuber, 2006; Hammann, Phan, Ehmer, & Grimm, 2008; Randler, Elker, Tempel, & Rehm, 2015). Students might therefore be overwhelmed if an experiment is too complex or they lack appropriate skills and knowledge (Arnold, 2015; Bruckermann et al., 2017; Schmidt-Weigand, Franke-Braun, & Hänze, 2008). If requirements exceed students' skills, they might perceive a high cognitive load (Kirschner, Sweller, & Clark, 2006), which can harm their motivation and learning (Ryan & Deci, 2017).

In this context, previous research frequently addressed cognitive aspects, such as the assessment of students' experiment-related skills and knowledge (e.g., Arnold, 2015; Chinn & Malhorta, 2002; Etkina et al., 2006; Hammann, Phan, & Bayrhuber, 2007; Hammann et al., 2006, 2008; Neumann, Schecker, & Theyßen, 2019). In addition to cognitive aspects, successful learning also depends on affective-motivational factors, such as intrinsic motivation (Ryan & Deci, 2017; Taylor et al., 2014) or flow experience (Engeser & Rheinberg, 2008). That is, student motivation must be considered when designing measures that optimally support students during learning. As a first step, variables that affect student motivation should be investigated to design such measures. In science education, this seems to be especially important for experimentation processes. As one of those variables, students' perceived competence was tested in this study in three phases of experimentation as a predictor of a specific situational motivation, namely, students' flow experience.

The need for competence refers to an individual's desire to feel effective in action and express and extend his or her skills (Reeve, 2015; Ryan & Deci, 2017). Individuals perceive competence if they perceive their skills and the challenge of their task as corresponding and are able to master the task (Deci & Ryan, 2000; Jang, Reeve, & Deci, 2010; Reeve, 2015; Ryan & Deci, 2017). The perception of competence can lead to intrinsic motivation (Reeve, 2015; Ryan & Deci, 2017). Intrinsically motivated actions are not driven by external incentives, but rather are an expression of interest and enjoyment (Ryan & Deci, 2017).

A complementary motivational quality to intrinsic motivation is flow (Deci & Ryan, 2000; Sheldon & Filak, 2008; Nakamura & Csikszentmihalyi, 2014; Reeve, 2015). *Flow* is a state in which individuals experience themselves as absorbed by their current action and perceive a smooth progression of their action (Nakamura & Csikszentmihalyi, 2014). During a flow experience, the individual perceives him or herself as optimally challenged (Nakamura & Csikszentmihalyi, 2014; Reeve, 2015; Taylor, Shepers, & Crous, 2006). That is, the challenge of an action is not perceived as too difficult or too easy, it corresponds to the perceived skills (*perceived balance between skills and challenges*; Nakamura & Csikszentmihalyi, 2014; Reeve, 2015).

Due to this common element of perceived competence and flow experience – namely, the perceived balance between skills and challenges (Deci & Ryan, 2000; Reeve, 2015) – a positive relationship between these variables is assumed, which has already been found in previous studies (e.g., Kowal & Fortier, 1999; Schüler, Sheldon, & Fröhlich, 2010; Taylor et al., 2006). Perceived competence is discussed as an antecedent of flow (Kowal & Fortier, 1999). However, the prediction of students' flow experience by their perceived competence during the phases of experimentation has thus far not been investigated in a differentiated manner. For this interrelationship, a closer look at the experimentation process reveals that a differentiated perspective can be useful.

As a problem-solving process, experimentation can be described using the scientific discovery as dual search model (SDDS) developed by Klahr and Dunbar (1988; see also Klahr, 2000). In SDDS, problem-solving occurs within and between two problem spaces, the *hypotheses space* and the *experiment space* (Klahr, 2000; Klahr & Dunbar, 1988). The search in these two spaces is described by three basic components: *Search Hypotheses Space*, *Test Hypothesis*, and *Evaluate Evidence* (Klahr, 2000; Klahr & Dunbar, 1988). The search for a hypothesis involves finding an appropriate research question and hypothesis within the hypotheses space (Search Hypotheses Space; Klahr, 2000; Klahr & Dunbar, 1988). To test this hypothesis, a meaningful and reasonably controlled experiment needs to be discovered in the experiment space and conducted afterwards (Test Hypothesis; Klahr, 2000; Klahr & Dunbar, 1988). The evaluation of evidence refers to the interpretation of the results of the current experiment and the consideration of further findings (Evaluate Evidence; Klahr, 2000; Klahr & Dunbar, 1988). The evaluation leads to a decision on the acceptance or (partial) rejection of the tested hypothesis (Klahr, 2000; Klahr & Dunbar, 1988).

Through the lens of SDDS, experimentation comprises an array of components and subcomponents of problem-solving processes. However, when experimenting in science class, not all of these (sub)components are usually considered (Seidel et al., 2002; Wirth, Thillmann, Künsting, Fischer, & Leutner, 2008). As teachers want to ensure that their students successfully perform the experimental tasks within a time-limited lesson, many teachers implement 'recipe-style' experiments in their science class (Abrahams & Millar, 2008; Abrahams & Reiss, 2012). For instance, the students work on a research question or hypothesis, conduct a pre-structured experiment following explicit instructions and analyse the collected data based on a problem given by the teacher. As this type of experiment is prevalent in school life (Abrahams & Millar, 2008; Abrahams & Reiss, 2012), we focused on three phases of experimentation: The development of a research question or hypothesis (Search Hypotheses Space), the conduct of an experiment (Test Hypothesis), and the analysis of the obtained results (Evaluate Evidence). In these three phases, students' perceived competence was investigated as a predictor of their flow experience.

RESEARCH QUESTION

Experiments offer the opportunity for hands-on activities in science class. However, students often lack skills and knowledge required for experimentation (Baur, 2018; Chinn & Malhorta, 2002; Etkina et al., 2006; Hammann et al., 2006, 2008; Randler et al., 2015). This gap can have a negative effect on students' motivation (Kirschner et al., 2006; Ryan & Deci, 2017). As motivation plays a key role in successful student learning (Engeser & Rheinberg, 2008; Ryan & Deci, 2017; Taylor et al., 2014), measures that optimally support students during experimentation should be developed. In this regard, the effects of motivational factors such as students' perceived competence during experimentation on their situational motivation must be investigated as a first step. In particular, the perception of competence during the three phases of experimentation should be separately considered. Therefore, we investigated the following research question:

Does students' perceived competence during the development of a research question or hypothesis (Search Hypotheses Space), the conduct of an experiment (Test Hypothesis), and the analysis of the obtained results (Evaluate Evidence) predict their flow experience?

METHODS

Sample

In this study, 212 German students (age: M = 16.30 years, SD = 0.96 years; 59% female) were taught about enzymology in biology class. The students attended the first year of upper secondary school (53% '*Gesamtschule*', i.e., comprehensive school; 47% '*Gymnasium*', i.e., high school).

Design and teaching content

A cross-sectional study was conducted in two subsequent lessons of 60 min each. At the beginning of the first lesson, all phases of experimentation described in the SDDS were explained to the students. Afterwards, the first experiment was performed. In the second lesson, the second experiment was performed. Immediately thereafter, students' perceived competence and their flow experience were assessed (Figure 1).

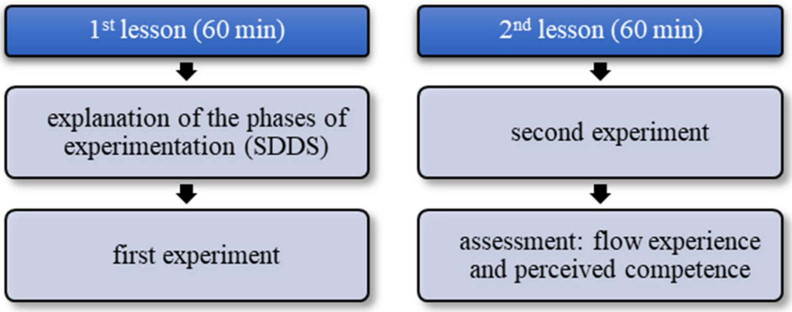


Figure 1. Study design.

The first experiment dealt with the catalysis of starch degradation by the enzyme amylase (α-amylase extracted from filamentous fungi *Aspergillus oryzae*). The students treated three 1% starch solutions stained with Lugol's iodine with amylase, saliva, and water. Students should have observed a decrease in the blue colouration of the mixtures containing amylase or saliva and no colour change in the mixture containing water. In that saliva contains amylase, the decrease in the colouration of these mixtures indicates that amylase catalyses the degradation of starch.

The second experiment regarded the temperature and pH dependence of the enzyme catalase (extracted from yeast *Saccharomyces cerevisiae*). In the first part of the experiment, the catalase was exposed to different temperatures. In the second part, the catalase was treated with different pH using hydrochloric acid (1%), water, and caustic soda (1%). Afterwards, hydrogen peroxide (10%) was added to the enzyme suspensions. Depending on the temperature and the pH, foam columns of different heights should have been formed in the test tubes, which were then correlated with the enzyme activity.

Measures

To assess students' perceived competence during experimentation, we used a self-developed test instrument. According to the three basic components of the SDDS (Klahr, 2000; Klahr & Dunbar, 1988), we adapted items of validated scales for the assessment of perceived competence (Van den Broeck, Vansteenkiste, De Witte, Soenens, & Lens 2010; Wilde, Bätz, Kovaleva, & Urhahne 2009). The test instrument consists of three subscales, each consisting of three items: *Search Hypotheses Space* ('While I was working on the hypotheses, I felt competent.'; Cronbach's Alpha: $\alpha = .74$), *Test Hypothesis* ('While I was conducting the experiments, I was convinced that I am able to do this properly.'; $\alpha = .71$), and *Evaluate Evidence* ('While I was analysing the data from the experiments, I felt that I was pretty good at it.'; $\alpha = .80$). A three-factorial confirmatory factor analysis with a maximum likelihood estimation with robust (Huber-White) standard errors (R-Studio with the lavaan package) yielded a good fit ($\chi^2(24) = 63.01$, $p < .001$, RMSEA = .08, SRMR = .05, CFI = .93; see Bryne, 2001; Hu & Bentler, 1999).

Students' flow experience was assessed using the flow short scale by Rheinberg, Vollmeyer, and Engeser (2003). The validated scale is a frequently applied test instrument (Engeser & Rheinberg, 2008; Stoll & Ufer, 2021). The flow short scale consists of 13 items ('I am completely absorbed in what I am doing.'; α = .70). The internal consistencies for both test instruments were satisfactory (see Field, 2018; Schmitt, 1996).

Statistics

The presumed relationships between the investigated variables were preliminarily investigated using Pearson correlation coefficients. Afterwards, a multiple regression analysis was performed to investigate the three aforementioned dimensions of students' perceived competence during experimentation as predictors of their flow experience.

RESULTS

The preliminary analysis revealed that the assumed predictors (*perceived competence during experimentation*) and the criterion (*flow experience*) were significantly correlated (Table 1).

Table 1. Descriptive statistics and Pearson correlation coefficients.

Variables		*M*	*SD*	1	2	3	4
Perceived competence during experimentation	Search Hypotheses Space	2.38	0.71	–	$.69^{***}$	$.77^{***}$	$.60^{***}$
	Test Hypothesis	2.62	0.68		–	$.67^{***}$	$.65^{***}$
	Evaluate Evidence	2.46	0.74			–	$.64^{***}$
Flow experience		2.54	0.46				–

Note: Variables range from 0 = strongly disagree to 4 = strongly agree; $N = 212$; *** indicates $p < .001$

The following regression analysis yielded a suitable model to investigate predictors of students' flow experience ($R^2 = .50$; $F(3, 208) = 69.39$, $p < .001$). Whereas students' perceived competence during the phase *Search Hypotheses Space* could not be confirmed as a predictor in this model, students' perceived competence during the phases *Test Hypothesis*, and *Evaluate Evidence* were predictors of their flow experience (Figure 2). The more competent the students perceived themselves during the test of the hypothesis and the evaluation of the evidence, the higher their flow experience during experimentation.

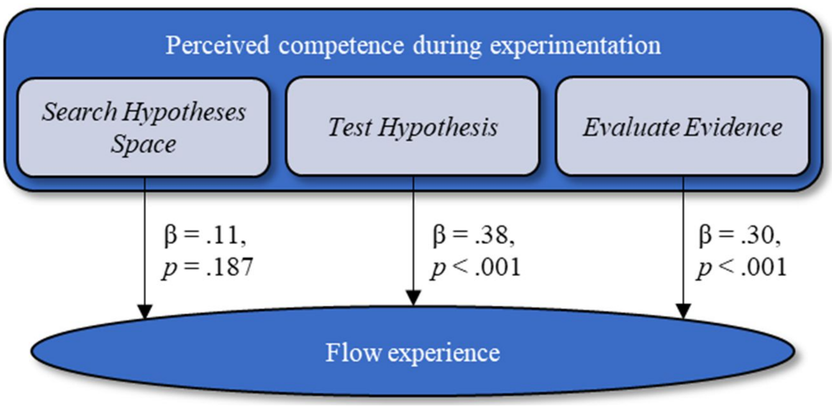


Figure 2. Results of the multiple regression analysis.

DISCUSSION

This study aimed to investigate the effects of students' perceived competence during experimentation on their flow experience. As previous studies found an impact of perceived competence on flow experience (Kowal & Fortier, 1999; Schüler et al., 2010; Taylor et al., 2006), we expected that students' perceived competence during the work on the hypothesis, the conduct of the experiment, and the analysis of the results would predict their flow experience. Our findings partially support this expectation.

Regarding the conduct of the experiments (Test Hypothesis) and the analysis of the results (Evaluate Evidence), students' flow experience was predicted by their perceived competence during these phases. That is, the more competent the students felt during the execution of the experiments and the analysis of the results, the more their flow experience was exhibited. Students' perceived competence during the conduct of the experiments was the strongest predictor. In contrast with the other phases, the test of the hypothesis mainly involves hands-on activities that might have led to a different quality of competence experience. Therefore, the perceived competence during this phase of the experimentation process might have been of special significance for the students' flow experience during experimentation.

However, students' perceived competence during the work on the hypotheses (Search Hypotheses Space) did not predict their flow experience. Students often do not make a connection between an experiment and the underlying hypothesis (Baur, 2018; Hammann et al., 2006, 2008; Hofstein & Lunetta, 2004; Randler et al., 2015). Therefore, the investigated students might not have been aware of the importance of their work on the hypothesis in the experimentation process. This could explain why students' perceived competence during the work on the hypotheses did not predict their flow experience.

Despite our interesting findings, we have to address some limitations. First, we used a self-developed test instrument to investigate students' perceived competence during experimentation. However, the scales consisted of adapted items derived from validated scales that assess perceived competence (Van den Broeck et al., 2010; Wilde et al., 2009); therefore, content validity is assumed to be sufficient. In addition, the internal consistencies were satisfactory (see Field, 2018; Schmitt, 1996). Furthermore, a confirmatory factor analysis revealed a good model fit supporting factorial validity (see Bryne, 2001; Hu & Bentler, 1999). To summarise, these indicators preliminarily suggest sufficient validity.

Second, some experimentation processes may involve additional or different components than those we investigated (see Klahr, 2000; Klahr & Dunbar, 1988). Therefore, it is conceivable to extend the test instrument with additional subscales that refer to further components of the SDDS, if needed. Moreover, students' flow experience might be affected by other student variables such as their perceived autonomy (Kowal & Fortier, 1999; Taylor et al., 2006). As teachers may pre-structure some phases of the experiments in their science class (Abrahams & Millar, 2008; Abrahams & Reiss, 2012), the students might feel controlled during these phases. This perceived lack of control would in turn impair their perception of autonomy and their experience of flow. Future studies should therefore include additional variables in the exploration of motivation during experimentation.

CONCLUSION

The results of this study contribute to a deeper understanding of motivational processes during experimentation. The perceived competence during various phases was found to be of varying significance for student motivation. This variance indicates that a differentiated investigation of the effects of perceived competence during experimentation might be worthwhile. In addition, our results hint that focus should be placed on perceived competence during the conduct of an experiment and analysis of the results to support student motivation. This suggested emphasis is of special importance for designing and evaluating competence-supportive measures during experimentation. Further studies might embrace this issue and design differentiated measures to support student motivation during experimentation.

ACKNOWLEDGEMENT

This project is part of the 'Qualitätsoffensive Lehrerbildung', a joint initiative of the Federal Government and the *Länder* which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research (funding code: 01JA1908). The authors are responsible for the content of this publication.

REFERENCES

Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education, 30*(14), 1945-1969. doi: 10.1080/09500690701749305

Abrahams, I., & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching, 49*(8), 1035-1055. doi: 10.1002/tea.21036

Baur, A. (2018). Fehler, Fehlkonzepte und spezifische Vorgehensweisen von Schülerinnen und Schülern beim Experimentieren [Mistakes, misconceptions, and pupils' idiosyncratic approaches to experimentation]. *Zeitschrift für Didaktik der Naturwissenschaften, 24*, 115-129. doi: 10.1007/s40573-018-0078-7

Bruckermann, T., Arnold, J., Kremer, K., & Schlüter, K. (2017). Forschendes Lernen in der Biologie [Scientific inquiry in biology]. In T. Bruckermann & K. Schlüter (Eds.), *Forschendes Lernen im Experimentierpraktikum Biologie* (pp. 11-26). Berlin, Germany: Springer. doi: 10.1007/978-3-662-53308-6

Byrne, B. M. (2001). *Structural equation modelling with AMOS. Basic concepts, application, and programming.* Mahwah, NJ: Erlbaum.

Chinn, C. A., & Malhorta, B. A. (2002). Children's responses to anomalous scientific data: how is conceptual change impeded? *Journal of Educational Psychology, 94*(2), 327-343.

Deci, E. L., & Ryan, R. M. (2000). The "What" and "Why" of goal pursuits. Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. doi: 10.1207/S15327965PLI1104_01

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion, 32*, 158-172. doi: 10.1007/s11031-008-9102-4

Etkina, E., Murthy, S., & Zou, X. (2006). Using introductory labs to engage students in experimental design. *American Journal of Physics, 74*(11), 979-986. doi: 10.1119/1.2238885

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). London, England: Sage edge.

Hammann, M., Phan, T. T. H., Ehmer, M., & Bayrhuber, H. (2006). Fehlerfreies Experimentieren. [Faultless experimentation]. *Der mathematische und naturwissenschaftliche Unterricht, 59*(5), 292-299.

Hammann, M., Phan, T. H., & Bayrhuber H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Dimensionen beim Experimentieren zu messen? [Experimentation as problem-solving: Can the SDDS-model be used to measure different dimensions of experimentation?]. *Zeitschrift für Erziehungswissenschaft, 10*(8), 33-49.

Hammann, M., Phan, T. T. H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education, 42*(2), 66-72. doi: 10.1080/00219266.2008.9656113

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science education, 88*(1), 28-54. doi: 10.1002/sce.10106

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling, 6*, 1-55. doi: 10.1080/10705519909540118

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588-600. doi: 10.1037/a0019682

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75-86. doi: 10.1207/s15326985ep4102_1

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1-48.

Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge, England: The MIT Press.

Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education, 33*(1), 27-50. doi: 10.1080/09500693.2010.518645

Kowal, J., & Fortier, M. S. (1999). Motivational determinants of flow: Contributions from self-determination theory. *The Journal of Social Psychology, 139*(3), 355-368. doi: 10.1080/00224549909598391

Nakamura, J., & Csikszentmihalyi, M. (2014). The concept of flow. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology* (pp. 239-263). Dodrecht, Netherlands: Springer. doi: 10.1007/978-94-017-9088-8_16

Neumann, K., Schecker, H., & Theyßen, H. (2019). Assessing complex patterns of student resources and behavior in the large scale. *The Annals of the American Academy of Political and Social Science, 683*(1), 233-249. doi: 10.1177/0002716219844963

Organisation for Economic Co-operation and Development (2016). *PISA 2015 results (volume I): Excellence and equity in education.* Paris, France: OECD Publishing. doi: 10.1787/9789264266490-en

Randler, C., Elker, S., Tempel, B. J., & Rehm, M. (2015). Missachtete Aspekte der Experimentierkompetenz [Disregarded aspects of the competence in experimentation]. *Der mathematische und naturwissenschaftliche Unterricht, 68*(5), 260-264.

Reeve, J. (2015). *Understanding motivation and emotion* (6th ed.). Hoboken, NJ: Wiley.

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens [The assessment of flow experience]. In J. Steinsmeyer-Pelster & F. Rheinberg (Eds.), *Diagnostik von Motivation und Selbstkompetenz* (Test und Trends Vol. 2, pp. 261-279). Göttingen, Germany: Hogrefe.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York, NY: The Guilford Press. doi: 10.1521/978.14625/28806

Schmidt-Weigand, F., Franke-Braun, G., & Hänze, M. (2008). Erhöhen gestufte Lernhilfen die Effektivität von Lösungsbeispielen? Eine Studie zur kooperativen Bearbeitung von Aufgaben in den Naturwissenschaften [The influence of different presentation modes of worked examples on learning]. *Unterrichtswissenschaft, 36*, 365-384.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4), 350-353. doi: 10.1037/1040-3590.8.4.350

Schüler, J., Sheldon, K. M., & Fröhlich S. M. (2010). Implicit need for achievement moderates the relationship between competence need satisfaction and subsequent motivation. *Journal of Research in Personality, 44*, 1-12. doi: 10.1016/j.jrp.2009.09.002

Seidel, T., Prenzel, M., Duit, R., Euler, M., Geiser, H., & Hoffman, L. (2002). „Jetzt bitte alle nach vorne schauen!" – Lehr-Lernskripts im Physikunterricht und damit verbundene Bedingungen für individuelle Lernprozesse [„Can everybody look to the front of the classroom please?" - Patterns of instruction in elementary physics classrooms and its implications for students' learning]. *Unterrichtswissenschaft, 30*, 52-77.

Sheldon, K. M., & Filak, V. (2008). Manipulating autonomy, competence, and relatedness support in a game-learning context. New evidence that all three needs matter. *British Journal of Social Psychology, 47*, 267-283. doi: 10.1348/014466607X238797

Stoll O., & Ufer, M. (2021). Flow in sports and exercise: A historical overview. In C. Pfeifer & S. Engeser (Eds.), *Advances in flow research* (2nd ed., pp. 351-376). Cham, Switzerland: Springer. doi: 10.1007/978-3-030-53468-4_13

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology, 39*, 342-358. doi: 10.1016/j.cedpsych.2014.08.002

Taylor, C. M., Schepers, J. M., & Crous, F. (2006). Locus of control in relation to flow. *Journal of Industrial Psychology, 32*(3), 63-71.

Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the work-related basic need satisfaction scale. *Journal of Occupational and Organizational Psychology, 83*, 981-1002. doi: 10.1348/096317909X481382

Wilde, M., Bätz, K. Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM) [Testing a short scale of intrinsic motivation]. *Zeitschrift für Didaktik der Naturwissenschaften, 15*, 31-45.

Wirth, J., Thillmann, H., Künsting, J., Fischer, H. E., & Leutner, D. (2008). Das Schülerexperiment im naturwissenschaftlichen Unterricht. Bedingungen der Lernförderlichkeit einer verbreiteten Lehrmethode aus instruktionspsychologischer Sicht [The student experiment in science classes. Conditions of the learning supportiveness of a widespread teaching method from an instructional psychology point of view]. *Zeitschrift für Pädagogik, 54*(3), 361-375. doi: 10.25656/01:4356

SPANISH STUDENTS' PERCEPTIONS OF THEIR SCHOOL SCIENCE: THE RELEVANCE OF SCIENCE EDUCATION (SECOND) PROJECT

María-Antonia Manassero-Mas[1] and Ángel Vázquez–Alonso[1]

[1]University of the Balearic Islands, Palma, Spain

This paper studies the Spanish students' perceptions of their school science classes and some additional comparisons across several variables in order to update affective data on science education from the students' voice. The theoretical and methodological background are drawn from the elaborations and contributions of the Relevance of Science Education Second (ROSES) project, from its predecessor, the Rose-2002 project, and from literature on science-related attitudes. The participants are 15-year-old Spanish students (n = 152), who answered the ROSE-2020 questionnaire. This study refers to the scale "my science class", which asks for students' agreement/disagreement on twelve items that depict several aspects of school science education. The findings report the aspects that conform the students' highest perceptions (interest, importance of science and curiosity) and their lowest perceptions (the intention to become scientists or to get a job in technology) on school science. Further, some comparisons across gender, science choice and along time pinpoint some significant differences across groups. The differences with previous Rose-2002 point out that both waves display the same overall profile for students' perceptions, yet the 2020 perceptions display progress across all items, making the overall picture of current science class more encouraging and optimistic than the previous one. However, the science career gap still remains a challenge for science education. Further, the students planning to choose science next year show significant higher agreements on science classes than the others, yet the gender differences seem widely fading. Beyond the positive achievements, developing students' interest and curiosity through science education and renewing the efforts to close the vocational and gender gap towards scientific and technological careers are discussed and some agencies recommended.

Keywords: perceptions of school science, gender issues, student choices

INTRODUCTION

The Relevance of Science Education (ROSE) project (thereafter Rose-2002) gave voice the students to diagnose their attitudes to science and technology worldwide twenty years ago. an overall pattern of disenchantment with STEM and many gender differences was reported for young people in Western countries, where the science career gap was dramatic, as few Western teens would like to become a scientist or a technologist, and extremely fewer girls (Sjøberg & Schreiner, 2019). Spanish teens fit somewhat the previous pattern, yet the gender differences were the lowest of Western countries (Vázquez & Manassero, 2007).

Recently, a new wave called ROSE Second (ROSES-2020) has been launched to update the evidence about young people's attitudes towards science worldwide. This paper aims to present some results of Spanish students' perceptions about school science classes and their gender, time and choice differences, and to discuss some recommendations for improving education within science, technology, engineering and maths (STEM) on the basis of extant evidence (Jidesjö et al., 2021).

Rationale and framework: 21st century challenges

The current knowledge societies are deeply embedded in STEM and many of the challenges they are facing are STEM-related challenges: environmental issues are perennial and COVID19 is the last but not least. International organizations like UNESCO, OECD and the European Union pay increasing attention to education, and in particular to STEM education, because STEM literacy plays key roles and gains importance in modern societies. STEM education is key both for the STEM-driven economies and for personal, social, cultural and democratic aims, where STEM-literate and participating citizenry is on the focus (OECD, 2016b).

Large-scale international achievement testing like TIMSS and PISA assess student cognitive learning and the related factors that may explain test scores. The affective factors are not the key concerns of TIMSS and PISA, as their main focus is cognitive (the knowledge of science) yet the last waves present some attention towards attitudes; for instance, PISA 2015 appreciated some increase in student enjoyment of science. Further, many top scoring countries on TIMSS and PISA tend to display low interest and attitudes to science, which is a main concern for STEM education (OECD, 2016a). The lack of perceived relevance for school science education is probably the greatest barrier to STEM literacy, to significant STEM learning and to develop interest toward STEM subjects, and these hindrances harm personal and social interests in the long run. Thus, positive attitudes towards STEM are worth and important learning goals for school education.

ROSES-2020 project gives again voice to students for gathering empirically-based insights into the affective contexts of STEM education to unveil the factors that conform their attitudes and motivations to learn STEM, thus complementing TIMSS and PISA cognitive aims. The focus on students' voice reflects the idea that students have something to say about their education and schools, the values, interests, and cultural backgrounds of individual and groups of students, and the corresponding pedagogies that are based on student choices, opinions, and ambitions. The literature on student voice develops the faithful attention to students' voice and their engagement and motivation in learning (Quaglia & Corso, 2014).

Rose-2002 study collected worldwide answers to "School Science Classes" (Sjøberg & Schreiner, 2019) and set up their general profile: school science was less interesting than other school subjects and displayed a strong gender difference pattern, with girls disagreeing more than boys in the wealthier countries. The statement claiming that science has opened my eyes for new and exciting jobs displays the same gender pattern and the lowest agreement scores for the richest countries. Three statements about school science (teaching to take care of my health, increasing my curiosity and showing the importance of STEM for our way of living) display less than half respondents agreeing with in most European countries.

Recently, Aschim et al. (2021) displayed the ROSES-2020 Norway results that represent a reference for this study on school science classes. The proportion of Norwegian students agreeing that school science is interesting is high (68%), yet in regard of school science opening eyes to new and exciting jobs the agreement rates stay around a low agreement rate (ca. 30%), slightly higher in ROSES-2020. The proportion of students that like school science better than most other subjects has increased from 33% (Rose-2002) to 41% (ROSES-2020). Further, the proportion of students agreeing that science will be useful in their everyday life shows a slight decrease (from 57% in Rose-2002 to 54% in ROSES-2020).

The picture of school science between Norwegian boys and girls are nuanced and interesting. In 2002 a significant proportion of boys (larger than girls) agreed that school science is interesting, better than most other subjects, helpful for everyday life, and opened to new and exciting jobs. In 2020 the gender differences have strongly decreased and the differences have even been reversed in some items, such as about school science has been helpful in my everyday life and opened eyes to new and exciting jobs (Aschim et al., 2021).

Impacts, aims and hypothesis

ROSES-2020 specifically aims to develop sound theoretical perspectives that are sensitive to the diversity of student backgrounds (cultural, social, gender, etc.) for evidence-based discussions related to education. As a consequence, ROSES-2020 aims to contribute recommendations for the improvement of STEM teaching, curricula, textbooks and classroom activities on the basis of the empirical findings about student voices and to raise some critical issues related to the relevance and importance of STEM and education for public debates on scientific and educational issues (i.e., reliable information, fake news, health, environment, etc.).

The innovation of ROSES-2020 project lies in the shift from mainstream cognitive knowledge (i.e., TIMSS and PISA) toward alternative criteria for success in education, which are related to innovative attitudinal aspects on STEM: interest, positive attitudes, willingness to engage in STEM issues, understanding the significance of STEM for our well-being and culture, etc.

This attitudinal shift in science education is closely linked with fostering sustainable, lasting and life-long results on STEM learning, through improving STEM literacy, motivation and engagement to STEM (school subjects, careers and occupations) and coping with gender differences. Overall, the results on gender differences are the most striking findings that discouragingly impact STEM learning and the recruitment to STEM. The low proportion of girls willing to choose STEM careers and occupations is a key concern in many countries. For instance, Spanish educational system at grades 11-12 splits up into vocational and pre-college studies, the latter enrolling a majority of girls (53.5%); then, pre-college students split again between science and non-science careers, where science careers enrol 52.7% girls, which is quite close of the global women enrolment. Thus, the Spanish statistics does not show any challenging gender gap between science and non-science studies at this overall level, yet some women underrepresentation appear at some STEM careers (physics, engineering or computers).

The ROSES-2020 data may provide equitable insights into how are configured the students' attitudes and opinions and how to increase girls' interest and motivation for STEM studies and careers. The overall hypothesis deal with the diagnostic of the current Spanish student attitudes and opinions about STEM education and its comparison to those diagnosed 20 years before. Further analysis across gender, type of school, choices, community, family books, etc. allow comparisons and hypothesis testing across groups and variables.

The research question posed here refers to the Spanish students' perception of school science classes: How do students perceive their school science? What features are perceived better and worse? These main questions are developed in regard of comparisons across gender, future choice and Rose-2002 previous results: How do ROSES-2020 results relate to the previous Rose-2002 results? Do the differences represent a progress or a regress? How do boys and girls perceive their school science classes? How does future science choice affect perception of school science classes?

METHODOLOGY

ROSES-2020 uses a mixed-method methodology, which involve qualitative and quantitative procedures to analyse the answers to the questionnaire and the comparisons between groups of students, yet for this preliminary study only quantitative data are used.

Participants

The ROSES-2020 target population are the students at the end of their compulsory education, because this educational stage allows students looking back to their education, reflect on what they have learned on STEM, and on how the educational choices for further studies are going to be made. This target population points out to students in late secondary education (aged 15) that corresponds in Spain to 3rd and 4th courses of the lower secondary education (grades 9-10). The participants in this study are 152 Spanish students (78 boys, 64 women and 10 blank), who validly answered ROSE-2020 questionnaire (average, 15.3 years). The participants are attending 4 publics and 2 public-funded private secondary schools, placed at a small city and at the suburb and the center of two different cities. The participants attend four public and two public-funded private secondary schools that are located at two small cities and at the suburb and the centre of two big cities, where they were surveyed by their teachers through whole group-class, which balances experiences across the independent variables (gender, choice, etc.)

Materials

The ROSES-2020 core methodological tool is a large questionnaire (ROSES-Q) that was developed by an international team of science education experts to gather the different affective data about the attitudes towards school and out-of-school experiences in science education, environment, interests, priorities, images and perceptions that are relevant for their STEM learning and literacy. Students answer ROSES-Q anonymously and they are free to let some items unanswered. ROSES-Q starts with a short presentation and a few background questions (sex, age, nationality and school and books at home); lastly, some country-specific items about city, region, and current engagement and future choice of STEM school subjects are asked.

The attitudinal categories of ROSES-Q involve interest in learning about STEM contents, priorities and motivations for a future occupation, views about environmental challenges, perceptions of school science classes and science education, perceptions on the role of STEM in society, use of ICT at home and school, out of school experiences related to STEM and two short open essays on “Myself as a scientist” and “the future occupation”.

The item wording is direct, simple, short, and avoids negative statements. Most items adopt the format of four-point Likert scale (1 – 4), where students are asked to tick the appropriate box number that best expresses their attitudes on the item. The meaning of box numbers varies across scales through the following categories: Disagree-Agree (the case for science classes scale), Not interested–Very interested, Not important–Very important, etc.

Procedures

The focus of this study is the scale of ROSES-Q categorized as “My science classes”, where students are asked to sincerely answer about their perceptions of 12 aspects of the science classes on a four-point Likert scale (table 1). The students’ responses on each item are codified as 1 (disagree), 2, 3 and 4 (agree); the agreement percent and weighted averages are computed to display simple and faithful representation of student opinions. Let’s take into account that item 1 has a negative wording (science is difficult) so that its score interpretation must be reversed: lower scores (disagreement) mean the students perceive science less difficult and higher scores (agreement) mean the students perceive science more difficult.

The data gathering took place from April until November 2020, when the schools were temporarily closed and the students were taught on line due to the COVID19 pandemic. Students anonymously and digitally answered the ROSES-2020 questionnaire as an on-line class assignment led and collected by their teachers at each participating school. The researchers and the secondary school teachers collaborated in real time to manage the administration process in each group class, the distribution to students the on-line links and the control of the reports on the questions and incidences during the surveying process.

RESULTS

The results are based on the participant students’ answers to the 12-item school science classes scale of the ROSES-Q. Table 1 summarises the descriptive statistics of students’ responses; the percent of agreement is computed by collapsing the students’ answers on the 3 or 4-point of the Likert scale. Further, the negative wording of item 1 that asks for the difficulty of school science subject requires a reversed interpretation of its scoring; for instance, 42.4% of agreement to difficulty means that a complementary majority of students (57.6%) do not perceive school science as difficult. Likewise, the average (2.40) must be reversed to its complementary (2.60), to join the same meaning of the remaining items.

The overall picture of student perceptions on school science is positive as most items (8) reach agreement rates over 50% and mean scores over the middle point of the Likert scale (2.5). However, some differences are appreciated across items; for instance, the top agreement items are item 2 (school science is interesting) and item 7 (school science has increased my curiosity), both showing large majority of students who perceive school science interesting (81.9%) and increasing the curiosity (70.7%).

On the other hand, the bottom agreement items are item 10 (wish to become a scientist) and item 11 (wish to get a job in technology), both reaching the lowest rates (around 30%) and the lowest average scores (around 1.9). These results depict the so-called vocational gap to STEM studies and careers among young people: few teens would like to enrol in STEM.

Table 1. Descriptive statistics for students’ responses to the items of the scale “My science classes”.

Item sentences	ROSES_2020				ROSES_2002	
	Mean (1-4)	St.D. Error	Standard deviation	Agree* (3-4)	Mean	SD
F1. School science is a difficult subject**	2.40	0.080	0.983	42.4%	2.58	0.77
F2. School science is interesting	3.26	0.070	0.867	81.9%	2.79	0.77

F3. School science has opened my eyes to new and exciting jobs	2.63	0.087	1.058	53.3%	2.28	0.89
F4. I like school science better than most other subjects	2.61	0.095	1.153	56.5%	2.31	1.00
F5. The things that I learn in science at school will be helpful in my everyday life	2.69	0.077	0.948	60.2%	2.46	0.82
F6. School science has made me more critical and sceptical	2.31	0.084	0.997	46.0%	2.07	0.81
F7. School science has increased my curiosity about things we cannot yet explain	3.01	0.085	1.033	70.7%	2.81	0.90
F8. School science has shown me the importance of science for our way of living	2.89	0.079	0.958	66.7%	2.65	0.82
F9. School science has taught me how to take better care of my health	2.55	0.083	1.008	56.5%	2.43	0.81
F10. I would like to become a scientist	2.03	0.091	1.106	31.2%	1.96	1.01
F11. I would like to get a job in technology	2.12	0.092	1.120	29.3%	1.90	0.99
F12. School science has helped me to understand sustainability solutions in my everyday life***	2.43	0.077	0.929	51.1%	-	-

* Percent of students who agree (scoring 3 or 4 on the Likert scale)
** This item displays a reverse writing (difficult); the interpretation of its scores must be inverted.
*** This item was not involved in Rose-2002

Comparison of the current ROSES-2020 results with the previous Rose-2002 results

The current sample data (ROSES-2020) are compared with the previous results (ROSE-2002) that were drawn from a larger sample at the same Spanish region (figure 1) just along the 11 items that share exactly the same content at both waves (item 12 of ROSES-2020 is excluded for not being included in Rose-2002).

The comparison depicts some relevant findings. First, the overall profiles of both waves are quite parallel (the apparent exception to parallelism of item 1 is due to the reversed meaning of item 1 for its negative formulation); thus, the student overall perception of science classes displays the same structure at both waves, which means that the highest, lowest and medium agreement items are roughly the same. For instance, item 2 (interesting) and item 7 (curiosity) have got the top scores, whilst student willingness to be a scientist or to have a job in technology have got the lowest scores at both waves, and the same applies for the intermediate items.

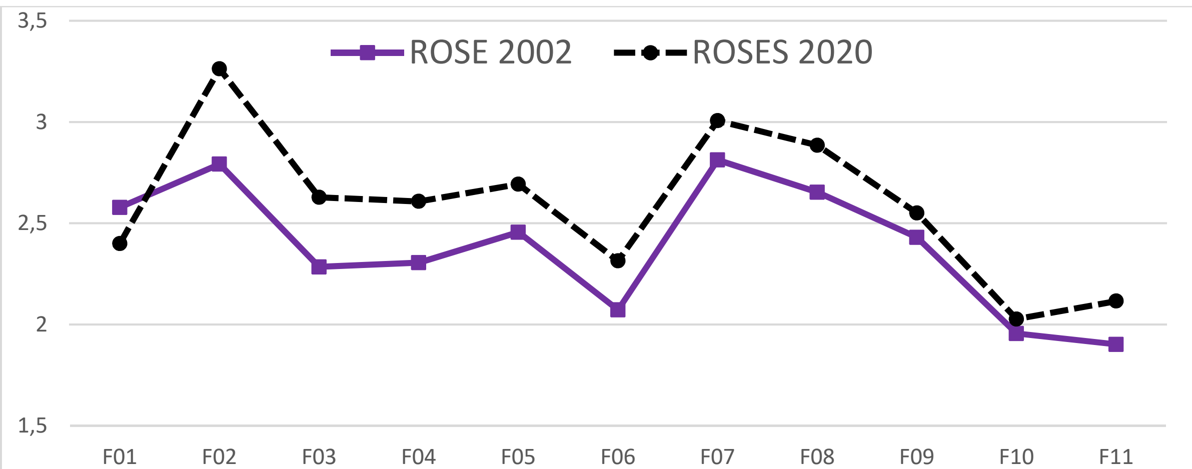


Figure 1. Average scores along the 11 common items of "My science classes" scale for Rose-2002 and ROSES-2020.

The most interesting finding is that ROSES-2020 perceptions display higher mean scores than Rose-2002 scores along all items, yet the effect size of the differences is only important (over half standard deviation; d = .57) for item 2 (school science is interesting). Thus, the overall meaning of this finding suggests that the current Spanish students hold a better perception of science classes than their mates' perceptions twenty years ago. Summing up, the student's appreciation of their science classes have improved along all items, and this global trend makes more relevant the progress of the students' perceptions.

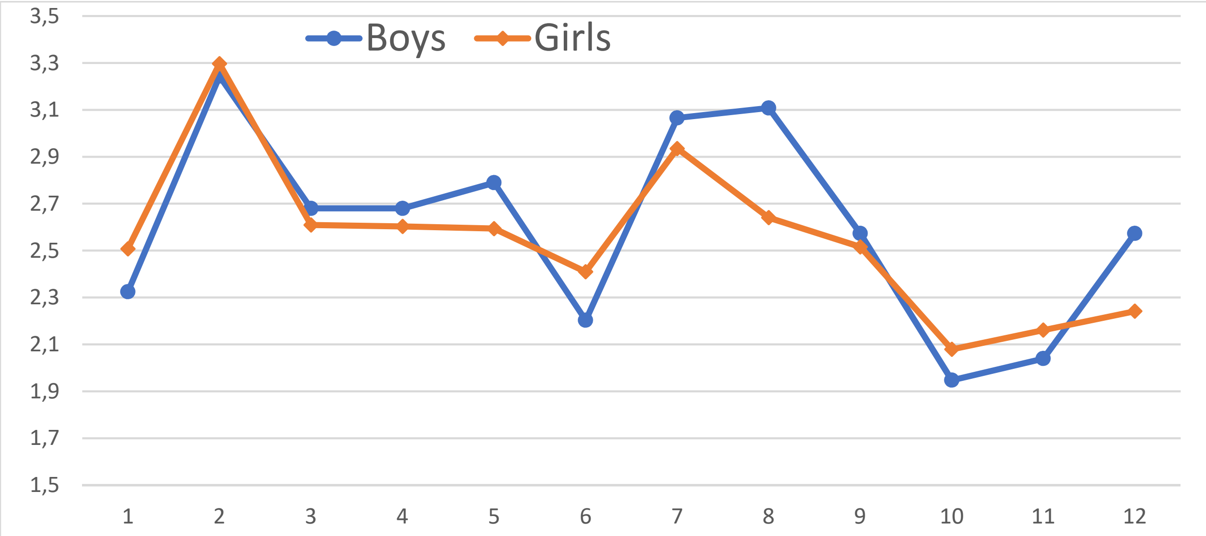


Figure 2. Average scores of "My science classes" scale items for boys and girls.

Comparison between boys and girls

The comparison between boys and girls along the items of the school science classes scale displays some significant findings (figure 2). Overall, boys' mean scores tend to be higher than girls' mean scores across a majority of items (8 out of 12 items); however, most differences are not statistically significant as only two items (8 and 12) display a moderate effect size for gender differences (d = .49 and d = .36 respectively).

Amid this expected outperformance of boys, the main finding to highlight on gender differences point out to girls higher scores than boys on their intention to be scientists (10) or to get a job in technology (11). The novelty of this result stems from the reversion of the usual gender stereotype on career gap (few girls hold a STEM vocation), as Spanish girls display stronger intentions than boys to pursue STEM careers, yet the differences are not significant.

Students expected future choice for science subject and perception of science classes

The organization of education in Spain compels 15-year-old Spanish students to make a choice between science and non-science subjects next year. Thus, ROSES-Q asked students whether they would choose a science subject, other subject or were undecided, and the perceptions of science classes have been analysed among these three response groups.

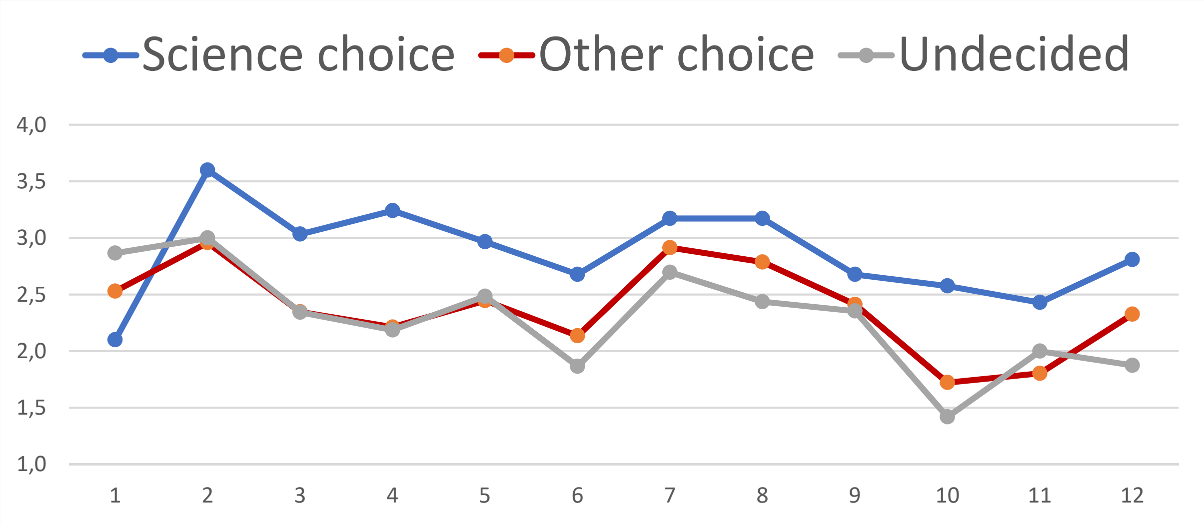


Figure 3. Average scores of "My science classes" scale items for the three student groups that next year will choose science, will make another choice or are undecided.

The results (figure 3) show that the students who will choose a science subject next year display the most positive perceptions of the science classes in regard of the two remaining groups (the undecided and the non-science choice groups), which in turn do not show significant differences between them. Further, half a dozen items of the perceptions on science classes (2, 3, 4, 5, 6 and 10) display relevant differences (effect size greater than half standard deviation) between the choosing science group and the others. This result suggests that improving the quality of classes may help the induction towards scientific careers and vocations.

CONCLUSIONS

The main finding of this study shows the overall student perceptions on school science classes are positive, as the majority of items display over 50% agreement and their weighted means are over the scale midpoint. Thus, the answer to the first research question is positive: Spanish students hold positive views about their school science classes. Further, the results allow pinpointing that the best perceived features of school science are its interest and its development of curiosity; on the other hand, the worst features correspond to the low rate of agreement to become scientists or to get a job in technology (the STEM vocational gap).

The comparisons with the baseline of Rose-2002 previous results confirm an overall similar profile of student perceptions of school science classes at both waves, which involves sharing the highest scored features (interest, importance of science and curiosity), as well as the lowest scored features (the intention to be scientists or to get a job in technology). This stability of profiles suggests some stability of "My science class" scale scores, which may lead to studying their psychometric properties (i.e., validity and reliability). Further, the main finding and positive news highlights that current scores are higher than Rose-2002 across all items, which points out that students' perceptions of school science classes have become more positive. This interpretation justifies the claim that the students' perceptions of school science show clear and positive progress in the last 20 years, which are also good news for Spanish STEM education. Thus, the general pattern of disenchantment suggested by Sjøberg and Schreiner (2019) does not apply to the perception of school science that stems from this Spanish sample.

Further, the gender differences still display the traditional stereotypical gendered pattern in favour of boys, in spite most differences are not statistically significant. Surprisingly, the most striking finding to highlight here is that girls outperform boys on their intention to be scientists or to get a job in technology, because it breaks the girls' traditional profile of lower vocations to STEM. These results are in line to the Norwegian students' overall reduction of gender differences on their perception of school science (Aschim et al. 2021).

The comparison of science classes perceptions across the students' intention to choose a science subject (or other subject) for the next year shows that the students who are willing to choose a science subject have got more significant and relevant perceptions of science classes than the remaining mates (undecided o willing to choose another subject). In spite this result may be expected, the good news is that science classes may induce students into the STEM pipeline.

In spite of the positive findings on science classes there is still room for improvement starting from some weakness reported here. An overall improvement aim must point out to increase all those features of science classes that have got the bottom scores; for instance, make students more critical and sceptical may be improved by developing their critical thinking skills that are also inherent to the scientific way of thinking. However, the main challenge would focus on the career gap to STEM studies, as it still remains a world concern of 21st century educational challenges for the current STEM-embedded societies (OECD, 2016a).

Finally, this study relates to the ESERA conference theme "Fostering scientific citizenship in an uncertain world" having in mind the severe social and educational changes resulting from the COVID19 public health problem, as it gives voice to students about their perception of school science in order to promote a student-centred and evidence-based framework specifically aimed to enhance STEM education (Quaglia & Corso, 2014). Taking the students' voices as the evidence to develop approaches to education that are aligned to the findings suggest the implementation of evidence-based pedagogy to teach science. Besides, the study also highlights the students as participating citizens, as their collaboration for assessing their school science classes is an exercise of students' responsible participation and engagement in their education (OECD, 2016b).

The main limitation of the present study arises from the tiny size of the sample, which makes the results tentative. As the project is currently in progress, it is expected that future larger samples will allow us to follow-up and, eventually, to confirm the trends presented here.

ACKNOWLEGMENTS

Grant PID2020-114191RB-I00 funded by Ministry of Science and Innovation/ 10.13039/501100011033.

REFERENCES

Aschim, E.L., Bjar, H. & Aae, R (2021). Changes in Norwegian students' interest in science, technology and environment. XIX International Organization for Science and Technology Education (IOSTE) Symposium - Transforming Science & Technology to Cultivate Participatory Citizens, Daegu, Korea 2020/2021. https://conf.ioste2020korea.kr/cms/index.php/ioste2020_op_229/

Jidesjö, A., Oskarsson, M., & Westman, A-K. (2021). Trends in Student's Interest in Science and Technology: Developments and Results from the Relevance of Science Education Second

(Roses) Study. XIX International Organization for Science and Technology Education (IOSTE) 2020 Symposium, Daegu, Korea. https://conf.ioste2020korea.kr/cms/index.php/2021/02/01/ioste2020_op_210/

OECD (2016a). *PISA 2015 Results (Volume I)*. PISA OECD Publishing.

OECD (2016b). *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. PISA OECD Publishing. http://dx.doi.org/10.1787/9789264267510-en

Quaglia, R., & Corso, M. (2014). *Student Voice: The instrument of change*. Corwin Press

Sjøberg, S., & Schreiner, C. (2019). *ROSE (The Relevance of Science Education) The development, key findings and impacts of an international low cost comparative project. ROSE Final Report, Part 1*. University of Oslo. https://www.academia.edu/40272545/The_ROSE_project._The_development_key_findings_and_impacts_of_an_international_low_cost_comparative_project_Final_Report_Part_1_of_2_

Vázquez, A., & Manassero, M.A. (2007). *La relevancia de la educación científica* [The relevance of science education]. Universitat de les Illes Balears.

CONTEXT-BASED LEARNING AS A METHOD FOR DIFFERENTIATED INSTRUCTION IN CHEMISTRY EDUCATION

Fabien Güth[1], Helena van Vorst[1]

[1]University of Duisburg-Essen, Essen, Germany

A well-known strategy to deal with heterogeneity in the classroom is called differentiated instruction. While the results of performance-based differentiation in chemistry education mostly do not show significant effects, research results on context-based learning indicate the potential of interest-based differentiation through contextual tasks with systematically varied characteristics. However, it is unclear which groups of students benefit from which contexts. The goal of this study is to identify groups of students that differ in terms of their context choice and to investigate how these students evaluate their choice after working on a context-based task in chemistry. For this purpose, a questionnaire study has been conducted with 349 third-year learners in chemistry in secondary schools. Through a cluster analysis based on individual student characteristics, four groups of students could be identified that show varying preferences at contexts with different characteristics. The results will be used for an interest-based differentiation with the help of systematically varied contextual tasks.

Keywords: Context-based Learning, differentiated instruction, Student Choices

THEORETICAL BACKGROUND

One of the expected impacts of the pandemic-related school closures is the increasing heterogeneity of students in class (Hammerstein et al., 2021). A central strategy to respond to this heterogeneity is called differentiated instruction. Differentiated instruction is an approach for teaching, in which the teacher takes the learning needs of small groups of students into account and varies his or her teaching accordingly (Tomlinson, 2016).

However, there is insufficient evidence for the effectiveness of differentiated instruction in secondary education (Smale-Jacobse et al., 2019). In chemistry education, first studies have been conducted that have investigated the effectiveness of differentiated instruction according to students' performance. Studies to date have mostly not found any effects of performance-based differentiation in chemistry education (e.g., Kallweit & Melle, 2014). As an additional approach, affective student characteristics (e.g., interest) could also be used to differentiate students' learning.

One way to address students' interests in chemistry is using out-of-school situations (so called contexts) with different overarching characteristics (Habig et al., 2018) . Bennett (2016) defines contexts as external situations that are used as a starting point for the development of subject matter. Furthermore, van Vorst et al. (2015) describe context characteristics, deduced from an in-depth literature review, to differentiate between contexts in science education systematically. One frequently observed context characteristic is *familiarity*. This characteristic is strongly dependent on the individual experiences of the students and is more closely described by the context characteristics *everyday* and *uncommon*. An *everyday* context has a strong relation to the living environment of the students (George & Lubben, 2002), while *uncommon* contexts deal with objects and situations that occur only rarely or not at all in the immediate students' daily life (Kasanda et al., 2005).

Habig et al. (2018) conducted a study to examine the effects of everyday contexts compared to uncommon contexts on student outcomes in chemistry education. The results indicate that students with a high level of interest and prior knowledge in chemistry benefit from uncommon contexts. In contrast, learners with a low interest and prior knowledge in chemistry benefit from everyday contexts. This leads to the assumption that individual student characteristics must be considered for an appropriate selection of contextual tasks.

For this reason, a further study investigated which contextual tasks are suitable for which students (van Vorst & Aydogmus, 2021). Students' context choice, their choice motives, and their satisfaction after completing the task in chemistry class were surveyed in this study. Three different clusters of students could be identified, which differed in terms of their context choice and choice motives. The first cluster consisted mainly of students with a low interest in chemistry. These students chose everyday contexts and indicated personal relevance as the most important motive for their choice. The second group summarized students with a high level of interest in chemistry. They chose uncommon contexts due to curiosity and interest. The third group was heterogeneous in terms of interest in chemistry and showed no clear contextual preference or choice motive. Additionally, these students were less satisfied with their chosen tasks compared to students from the first and the second cluster.

Based on these results, van Vorst & Aydogmus (2021) assumed that the third cluster may include students who are more interested in chemistry itself and do not need a context for their learning. A task without context uses a chemical situation (e.g., a situation from the laboratory) instead of an external situation as a starting point for learning. Based on the state of research to date, it is still unclear how to characterize students who choose everyday contexts, uncommon contexts, or tasks without a context. However, this information is necessary to offer appropriate contexts to different students in interest-based differentiated instruction.

RESEARCH QUESTIONS

The main goal of the study presented in this paper is to investigate which students choose which contextual tasks (everyday context, uncommon context, or non-contextualized task). Moreover, the aim is to examine how students evaluate their choice decision after working on their chosen task to find out whether the task suited the students. The following research questions will be answered.

(1) Which groups of students can be distinguished in their context choice with regard to their personal characteristics and which contextual preference do they show?

(2) How do students evaluate their context choice in terms of satisfaction, situational interest, and cognitive load after working on a chosen context-based task?

METHOD

To answer the research questions, a quantitative study with the help of questionnaires has been conducted. The context choice of 349 students (male: 49.9%; female: 50.1%) of the third year in learning chemistry from seven different secondary schools in Germany has been analyzed.

The study has been conducted using tablet computers in regular chemistry classes within 90 minutes. Figure 1 gives an overview of the proceeding of the study.

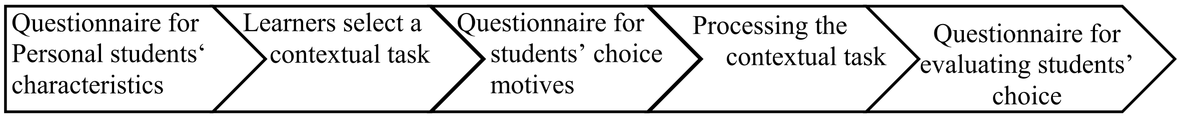


Figure 3. Design of the study.

In a first step, students’ individual variables like their demographic data, reading comprehension (Schneider et al., 2017), content knowledge in chemistry (Celik & Walpuski, 2018), individual interest and motivation in chemistry (Wild & Krapp, 1995), their chemical self-concept (Hoffmann et al., 1998), the leisure interest (Albert et al., 2019) as well as their career choice perspective (Kunter et al., 2002) have been surveyed. Afterwards, five different tasks have been presented to the students: two tasks were dealing with an everyday context, two tasks were dealing with an uncommon context and the last task was dealing with the chemical content without a context (Tab. 1).

Table 5. Developed contextual tasks and their characteristic affiliation.

Context characteristic	Context
Relation to everyday Life	Why brushing teeth is so important
	Our digestion
Uncommon Phenomena	The consequences of chronic gastritis type a
	Bone damage due to chronic renal insufficiency
Non-contextualized	Acids in the laboratory

To ensure that the context choice can be clearly attributed to the corresponding context characteristic, all tasks were completely identical except for the underlying context characteristic. All tasks addressed the chemical content acidic and alkaline solutions, which is rather interesting for most students (Habig, 2017). This was to ensure that enough students chose the non-contextualized task to characterize this group in more detail. Furthermore, all contextualized tasks focused on a context from the field of the human body and human diseases to address the interests of both genders as far as possible (Sjøberg & Schreiner, 2010). Lastly, all tasks were identical with regard to task formulation and external characteristics (e.g., layout, text structure, text length). In the next step, students’ choice motives (van Vorst & Aydogmus, 2021) have been investigated by using another questionnaire. Afterwards, students worked on their selected task. Finally, a questionnaire on satisfaction (own development), situational interest (Engeln, 2004), and cognitive load (Schwamborn et al., 2011) have been used to evaluate the choice decision.

RESULTS

Statistical data were analyzed using the statistical software *R*. Different packages were used, which significantly extend the functionality of R. We examined the quality of the knowledge test using Item Response Theory (Bond et al., 2021). Exploratory factor analyses and reliability analyses were used to examine the quality of the affective test instruments. Significant differences in person characteristics were analyzed using a multivariate analysis of variance

(MANOVA) followed by multiple analyses of variance (ANOVAs) to avoid the inflation of the familywise error rate (Field et al., 2013). Subsequently, we conducted a K-means cluster analysis to identify homogeneous groups of students. The number of clusters was determined using different measures such as the elbow criterion or the silhouette coefficient (Everitt et al., 2011). Linear regression analyses were performed to examine relations (Field et al., 2013).

Quality of the survey instruments

The quality of the knowledge test has been analyzed using *Item Response Theory*. We analyzed an one-dimensional Rasch model (e.g., Bond et al., 2021). The analysis of the model fit showes that all test items fit the calculated Rasch model (0.87 ≤ wMNSQ ≤ 1.10; -1.99 ≤ ZSTD ≤ 1.55). Despite this, the test indicates a questionable reliability (WLE-Reliability = .63).

For the analysis of the other test instruments, we conducted an exploratory factor analyses with subsequent reliability analyses (Field et al., 2013). The scales formed by factor analyses show sufficient reliabilities (.68 ≤ Cronbach's α ≤ .94).

Students' context choice

A majority of the students (63.6%) in this sample have chosen an everyday context. Furthermore, 20.1% have chosen an uncommon context and 16.3% have chosen the non-contextualized task.

To respond to the first research question, students have been grouped based on their personal characteristics. For this, variables that contribute to the separation of the groups have to be selected first. We have examined the personal characteristics that distinguish the students who have chosen a particular contextual task. The personal characteristics of the students who selected a certain contextual characteristic were compared for differences descriptively and inferentially (using a MANOVA and multiple ANOVAs). There are statistically significant differences with regard to performance ($F(2, 346) = 15.99$, $p < .001$, $\eta^2 = .08$), individual interest ($F(2, 346) = 5.80$, $p < .01$, $\eta^2 = .03$), chemistry-related self-concept ($F(2, 346) = 16.55$, $p < .001$, $\eta^2 = .09$), the choice motive of personal relevance ($F(2, 344) = 100.88$, $p < .001$, $\eta^2 = .37$), and the choice motive of surprising information ($F(2, 344) = 15.99$, $p < .001$, $\eta^2 = .08$). A one-way MANOVA shows a statistically significant difference between the students who have chosen a particular contextual task on the combined dependent variables ($F(2, 682) = 27.1$, $p < .001$, partial $\eta^2 = .27$, Wilk's $\Lambda = .51$).

A K-means cluster analysis has been conducted to identify groups of students who are as similar as possible regarding these personal characteristics. Analysis point to a four-cluster solution. Figure 2 shows the mean scores for the four clusters of students' personal characteristics.

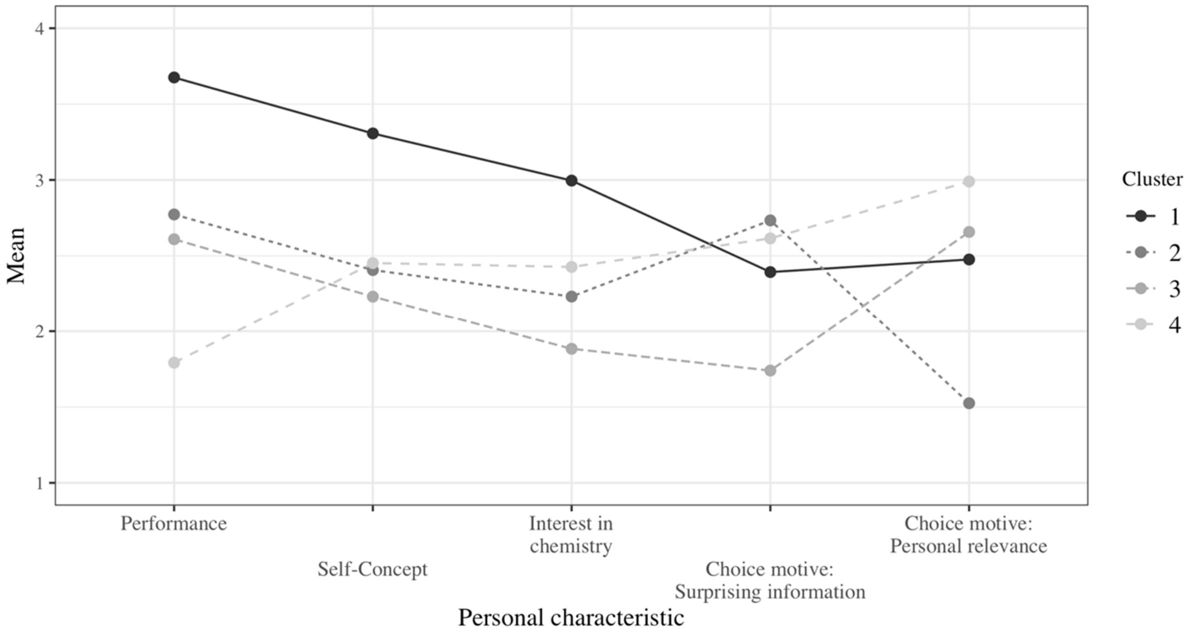


Figure 4. Mean Scores for the four clusters of students' personal characteristics. The person abilities from the Rasch model were transformed to a range of values from 1 to 4 for clarity.

Students from the first cluster show the highest performance, the highest self-concept, and the highest interest in chemistry in this sample. The choice motives surprising information and personal relevance are almost equally important in this cluster. 37.8% of these students have chosen the non-contextualized task, which is about twice as often as in the whole sample. Nevertheless, more students (48.8%) have chosen the everyday context.

The second cluster includes students with a lower performance, self-concept, and interest in chemistry. The primary choice motive is surprising information. Students from this cluster have chosen uncommon contexts to 60.0%. This is almost three times higher compared to the total sample.

Students with a lower performance, interest and self-concept compared to the students from the second cluster make up the third cluster. The primary choice motive is personal relevance. 84.5% of the students have chosen an everyday context. This is almost 20% higher than in the whole sample.

The last cluster contains students with a similar self-concept and a slightly higher interest compared to students from the second cluster. However, students in the fourth cluster show a much lower performance. The choice motive surprising information is almost as important as in the second cluster, but the personal relevance of a context is an even more important motive for choosing a context. This leads to the fact that 81.1 % of the students in the fourth cluster have chosen an everyday context.

Furthermore, we examined to what extent the clusters differ with regard to reading comprehension. Our results show that students from the last cluster (M = 17.51, SD = 12.94) have a much lower reading comprehension compared to students from the first (M = 26.30, SD = 12.94), second (M = 25.40, SD = 12.78), or third (M = 22.62, SD = 12.78) cluster. An ANOVA points to a statistically significant difference (F (3,330) = 7.65, p < .001, η^2 = .07). The Tukey post-hoc test indicates significant differences at $p < .05$ between the last and the other clusters, with no significant difference between the first three clusters.

Evaluation of context choice

The descriptive and inferential statistical analysis show that students in a certain cluster do not differ in satisfaction, situational interest or cognitive load depending on the chosen contextual task.

Satisfaction and situational interest after working on the selected task mainly depend on personal characteristics such as interest in chemistry. Results of a linear regression analysis show that interest in chemistry significantly predicts situational interest ($F(1,333) = 82.43$, $p < .001$, $R^2 = .20$) after working on the chosen task.

DISCUSSION

The results show that most learners choose an everyday context for learning chemistry. Other studies also found a preference for contexts with a personal relevance (Broman et al., 2020).

The context choice from the first cluster indicates that learners with high performance and high interest in chemistry are more likely to choose non-contextualized tasks. This cluster could be a group of learners that has also been identified in the study by van Vorst & Aydogmus (2021). This group of students might not be interested in contexts and thus, is not influenced by any context. The authors hypothesized that this type of student may be high-performing and/or highly interested, and not in need of an additional context. Based on our data, it can be shown that this type of students choose non-contextualized contexts more often. It is possible that a context is perceived as disruptive by this group because they want to learn more about chemistry itself. Only hypotheses can be made about the exact reasons since no clear choice motive could be identified in this cluster. However, this is also because the test instrument for assessing the choice motives did not include a scale for measuring the motives for choosing a non-contextualized task.

The second cluster indicates a type of students who is interested in chemistry and shows an intermediate level of performance. In addition, these learners choose a context based on surprising information. This is also reflected in their choice of a context, as a large proportion of these students choose the uncommon context. A higher situational interest of highly interested students in uncommon contexts has also been found by Habig et al. (2018). The authors conclude that the uncommon context is a special incentive for highly interested people to expand their knowledge.

Learners with the lowest performance, interest and self-concept belong to the third cluster. The primary motive for choice in this group of learners is personal relevance. Accordingly, more than 80% of the learners also choose an everyday context. Other studies also show a preference of low-interest students for contexts with a personal relevance (Habig et al., 2018; van Vorst & Aydogmus, 2021). We conclude that a connection to everyday life is probably needed for uninterested and low-performing students to engage in chemistry at all.

The profile of the fourth cluster is rather unusual, with learners showing the lowest performance but medium interest and self-concept. In comparison with the other clusters, it is noticeable that the reading comprehension of this group of learners is significantly lower than that of the other groups. For this reason, it is doubtful to what extent the context choice can be interpreted, as the learners may not have understood the description of the context.

After the learners' context selection and task processing, no differences in the clusters could be identified in terms of satisfaction, situational interest, and cognitive load. Therefore, we assume that all learners have chosen an appropriate context for themselves and hence, were satisfied with their choice after task processing.

OUTLOOK

The results could be used for data-supported differentiation in chemistry education (c.f., learning analytics; Clow, 2013) . A further study will be conducted to investigate the effects of an interest-based differentiated instruction, using contexts with different characteristics in chemistry. For this purpose, a recommendation system (Khanal et al., 2020) will be developed that suggests a context to students based on their individual characteristics.

REFERENCES

Albert, M., Hurrelmann, K., & Quenzel, G. (2019). *Jugend 2019: eine Generation meldet sich zu Wort [Youth 2019: a generation speaks out]*. Beltz.

Bennett, J. (2016). Bringing Science to Life. In A. Pilot, P. d. Brok, & R. Taconis (Eds.), *Teachers creating context-based learning environments in science* (21-39). SensePublishers.

Bond, T., Yan, Z., & Heene, M. (2021). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Routledge.

Broman, K., Bernholt, S., & Christensson, C. (2020). Relevant or interesting according to upper secondary students? Affective aspects of context-based chemistry problems. *Research in Science & Technological Education*, 1–21. https://doi.org/10.1080/02635143.2020.1824177

Celik, K., & Walpuski, M. (2018). Development and Validation of Learning Progressions on Chemical Concepts. In O. Finlayson, E. McLoughlin, S. Erduran, & P. Childs (Eds.), *Electronic Proceedings of the ESERA 2017 Conference. Research, Practice and Collaboration in Science Education* (pp. 1423–1432). Dublin City University.

Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, *18*(6), 683–695. https://doi.org/10.1080/13562517.2013.827653

Engeln, K. (2004). *Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken [Student laboratories: authentic, activating learning environments to spark interest in science and technology]*. Logos Verlag.

Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. *Wiley series in probability and statistics*. Wiley. https://doi.org/10.1002/9780470977811

Field, A., Miles, J., & Field, Z. (2013). *Discovering statistics using R*. Sage.

George, J. M., & Lubben, F. (2002). Facilitating teachers' professional growth through their involvement in creating context-based materials in science. *International Journal of Educational Development*, *22*(6), 659–672. https://doi.org/10.1016/S0738-0593(01)00033-5

Habig, S. (2017). *Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren [Systematically varied contextual tasks and their influence on cognitive and affective student factors]*. Logos Verlag.

Habig, S., van Vorst, H., & Sumfleth, E. (2018). Merkmale kontextualisierter Lernaufgaben und ihre Wirkung auf das situationale Interesse und die Lernleistung von Schülerinnen und Schülern

[Characteristics of contextualized learning tasks and their effect on students' situational interest and learning performance]. *Zeitschrift Für Didaktik Der Naturwissenschaften*, *24*(1), 99–114. https://doi.org/10.1007/s40573-018-0077-8

Hammerstein, S., König, C., Dreisoerner, T., & Frey, A. (2021). *Effects of COVID-19-Related School Closures on Student Achievement - A Systematic Review.* https://doi.org/10.31234/osf.io/mcnvk

Hoffmann, L., Häußler, P., & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik [The IPN Physics Interest Study]*. IPN.

Kallweit, I., & Melle, I. (2014). Effects of self-evaluation on students' achievements in chemistry education. In C. P. Constantinou, N. Papadouris, & A. Hadjigeorgiou (Eds.), *Science Education Research For Evidence-based Teaching and Coherence in Learning. Proceedings of the ESERA 2013 Conference.* European Science Education Research Association.

Kasanda, C., Lubben, F., Gaoseb, N., Kandjeo-Marenga, U., Kapenda, H., & Campbell, B. (2005). The Role of Everyday Contexts in Learner-centred Teaching: The practice in Namibian secondary schools. *International Journal of Science Education*, *27*(15), 1805–1823.

Khanal, S. S., Prasad, P., Alsadoon, A., & Maag, A. (2020). A systematic review: machine learning based recommendation systems for e-learning. *Education and Information Technologies*, *25*(4), 2635–2664. https://doi.org/10.1007/s10639-019-10063-9

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente [PISA 2000: Documentation of the Survey Instruments]*. Max-Planck-Institut für Bildungsforschung. http://hdl.handle.net/hdl:11858/00-001M-0000-0023-9987-C

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5-12: Lesegeschwindigkeits- und -verständnistest für die Klassen 5-12: Manual [LGVT 5-12: Reading speed and comprehension test for grades 5-12: Manual]*. Hogrefe.

Schwamborn, A., Thillmann, H., Opfermann, M., & Leutner, D. (2011). Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior*, *27*(1), 89–93. https://doi.org/10.1016/j.chb.2010.05.028

Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence. *Frontiers in Psychology*, *10*, 2366. https://doi.org/10.3389/fpsyg.2019.02366

Tomlinson, C. A. (2016). *The differentiated classroom: Responding to the needs of all learners*. Pearson.

van Vorst, H., & Aydogmus, H. (2021). One context fits all? – analysing students' context choice and their reasons for choosing a context-based task in chemistry education. *International Journal of Science Education*, 1–23. https://doi.org/10.1080/09500693.2021.1908640

Wild, K.-P., & Krapp, A. (1995). Elternhaus und intrinsische Lernmotivation [Parental home and intrinsic motivation to learn]. *Zeitschrift Für Pädagogik*, *41*(4), 579–595.

TEACHING STRATEGIES TO PROMOTE METACOGNITIVE REGULATION IN SECOND LEVEL CHEMISTRY EDUCATION

James Trimble[1], Odilla E. Finlayson[1] and James Lovatt[2]

[1]School of Chemical Sciences & Centre for Advancement of STEM Teaching and Learning, Dublin City University, Dublin, Ireland

[2] School of STEM Education, Innovation & Global Studies & Centre for Advancement of STEM Teaching and Learning, Dublin City University, Dublin, Ireland

This study aimed to address a gap between theory and practice in relation to students' metacognitive regulation in the upper second-level chemistry classroom. The gap pertains to the theoretical evidence that metacognitive regulation is an important skill in learning chemistry whereas attempts to incorporate specific teaching strategies in authentic classrooms are rare in comparison to traditional teaching approaches. To do so, metacognitive teaching and learning strategies were implemented with an upper second level class group studying chemistry over the course of an academic year, which included explicit metacognitive teaching, metacognitive modelling by teacher and student, think aloud protocol, regulatory checklists and pre and post lesson reflections. All lessons were audio and video recorded. A rubric was developed to analyse the teacher actions and dialogue in relation to metacognitive regulation. Students completed the Metacognitive Awareness Inventory at the beginning and end of the academic year. A significant increase was identified in students' overall metacognitive regulation, as well as in specific subcategories thereof.

Keywords: Metacognition, Teaching Practices, Conceptual Change

INTRODUCTION

Georghiades (2004) identified a theory-practice gap – which still exists today – where academic studies highlight the value of metacognition for science learning but reported attempts to highlight metacognition in classrooms are rare (e.g., Davidson et al, 1998). In agreement with Georghiades (2004), Thomas (2012) argues that while there are few researchers who question the importance of metacognition, the recognition of this importance is not reflected in teachers' or teacher educators' practices. Thomas (2012) also stresses that the extent to which teachers themselves are metacognitive is not clear and calls for more research on teacher metacognition because it might enable increased effectiveness of professional development in this area. Metacognitive regulation involves controlling one's cognitive processes and refers to the management of cognitive activities during learning, e.g., making changes in processes or strategies as a result of monitoring (Flavell et al., 2002; Whitebread et al., 2009). Though no universally accepted definition of metacognition can be found in the literature (Dinsmore, Alexander, & Loughlin, 2008; Harrison & Vallin, 2018), a longstanding view is that it comprises two components: knowledge of cognition and regulation of cognition. How metacognitive regulation can be improved through the implementation of teaching and learning strategies which address it are at the centre of the current study. Previous research has taken place into a range of teaching and learning strategies which highlight student metacognitive regulation (e.g. Schön, 1987; King, 1998; Tien, Rickey, & Stacy, 1999; Thomas & McRobbie, 2001; Chiu, & Linn, 2012; Ellis & Denton, 2013; Trimble, Lovatt & Finlayson, 2019).

Metacognitive regulation in this research refers to the regulation or control of one's cognition and one's knowledge about planning, implementing, evaluating, and correcting behaviours. The definition of the subcategories of the metacognitive regulation adopted in this research are outlined in Table 1.

Table 1. Subcategories of the metacognitive regulation

Metacognitive regulation		
Sub category	**Description**	**Example(s)**
Planning	Setting goals, selecting appropriate strategies, making predictions, strategy sequencing and allocating resources.	When one formulates the goal of an assignment or problem, and also decides which steps are needed to get the answer.
Monitoring	Monitoring of one's own cognitive processes which involves awareness and assessment of comprehension and task performance and progress towards the desired goal	In relation to a calculation, one checks whether every step has been calculated well and whether one has made any careless mistakes.
Evaluation	Entails an assessment of the products and efficiency of one's learning and thinking, for example through self-checking, reflection and re-evaluation, reviewing task performance, learning processes.	In relation to a calculation, one compares the outcome of the calculation to one's initial estimate, or do a recalculation.

When engaging in higher order thinking, students need to undergo specific metacognitive skills like monitoring their thinking process, checking whether progress is being made toward an appropriate goal, ensuring accuracy, and making decisions about the use of time and mental effort (Halpern, 1998). In light of this, the current research investigates the incorporation and implementation of teaching and learning strategies which specifically address student metacognitive regulation with a second level chemistry class group. The specific research questions addressed were:

1. *Are teaching and learning strategies which focus on developing student's metacognitive regulation identified in the literature suitable for integration into upper second level chemistry lessons?*
2. *How are metacognitive teaching and learning strategies which address metacognitive regulation evaluated in upper second level chemistry lessons?*
3. *Is there a significant difference between students' pre-test and post test scores on measures of student metacognitive regulation?*

In the initial stages of this research, teaching and learning strategies were identified in the literature which encourage metacognitive regulation in students (Schon, 1987; Schunk, 1989; King, 1991; Stensvold and Wilson, 1992; Butler & Winne, 1995; Elder and Paul 1998; Schraw, 1998; Rickey and Stacey 2000; Thomas and McRobbie, 2001; Gilbert, 2005; Kipins and Hofstein 2007). Following this, two pilot studies were carried out with students who were completing a module in chemistry before entering upper second level education. An analysis of the strategies implemented in these pilot studies identified a number of teaching and learning strategies which are suited for upper second level chemistry lessons (as summarised in table 2). A metacognitive teaching plan was developed by the researcher-practitioner to this end. It was piloted in the academic year 2018-19 and was implemented throughout the academic year 2019 –2020. Implementation of the teaching plan was evaluated through analysis of teacher actions and dialogue, class discussions, classroom activities, student generated work as well as a metacognitive student journal which was developed as part of this research. A rubric for analysing teacher actions and dialogue which promote metacognitive regulation was also developed.

Assessment and evaluation of students' metacognition regulation presents a problem, due to the nature of observing and/or quantifying a phenomenon which relates to one's own thoughts about how aware of, or in control of, one is of one's thoughts. Previous researchers (Winne & Perry, 2000; Muis et al. 2007; Pieschl et al. 2012) have argued metacognition can be assessed as either an aptitude or an event. When metacognition is viewed as an aptitude, students are assumed to possess tactics of metacognition as a relatively enduring trait or to employ relatively stable approaches when solving different tasks When metacognition is viewed as an event, the utilisation of metacognitive tactics is assumed to vary dramatically across different tasks/events, and students are assumed to adjust their learning behaviour dynamically to adapt to different context or task/event demands. Pintrich et al. (2000) argue that including metacognitive regulation, is similar to other kinds of knowledge stored in the long-term memory and that it can be accessed when properly cued. In this line of reasoning, they argue that self-reports are appropriate, as an easy and efficient measurement. The Metacognitive Awareness Inventory (MAI) is a self-report instrument designed to assess general self-regulated learning skills across the disciplines. Developed by Schraw and Dennison (1994), The MAI was developed specifically to address the two theoretical components (or dimensions) of metacognition: knowledge of cognition (17 items) and regulation of cognition (35 items).

Measuring metacognition as an event requires examining the metacognitive skills in real time to capture the dynamic processes within a particular task. Methods are required that allow researchers to document, identify and examine the target behaviours and verbalisations as they occur in real time during an authentic event or situation (Cleary, 2011).

METHOD

Participants

The study was concerned with the teaching and learning of upper second level chemistry in authentic upper second level chemistry lessons. The class group of students in this research were from a non-fee-paying school, drawing from across academic and socioeconomic backgrounds. Data was collected before (MAI), during (video and audio recording, student generated work) and after (MAI) the implementation of the teaching plan. The student participants numbered 24 male students aged between 17-18 years old. A convenient sample was used in this research as the participants are part of the researcher/practitioner's normal class group who have chosen to study chemistry and were in year one of their final two years of upper second level education.

Teaching and Learning Strategies

The teaching and learning strategies implemented and evaluated in lessons which addressed promoting student metacognitive regulation are outlined in Table 2. Each of the activities were developed in the pilot studies of this research and were included in the lessons throughout the academic year in the current study. The sequencing of topics was such that initial topics (e.g., states of matter, gases laws, ionic and covalent bonding, acids and bases etc) where addressed through the lens of concepts such as electronegativity, electrostatic attraction, collision theory, bond energy and consequently building on these threshold concepts. The aim of order of topics was to provide a conceptual structure in which students could improve metacognitive regulation. A student metacognitive chemistry journal contained student templates for activities which promoted metacognitive regulation e.g., pre and post lesson reflections; visualisation of concepts sheets; regulatory checklists; homework and test wrappers.

Table 2. Teaching and learning strategies to promote and incorporate metacognitive regulation in lessons.

Strategy	Details
Metacognitive prompts and questioning;	Metacognitive cues, questions or regulatory checklists that were used by the students during activities such as problem-solving, experimentation, inquiry learning, reading chemistry texts, writing reports and reflections, or discussing science topics
Reflective writing;	Students were asked to write a metacognitive journal about their thinking, visualisation and learning, pre and post lesson surveys, or written responses to reflective prompts.
Visualisation of concepts	Students were asked to link visualisation of concepts in chemistry with the development of their understanding of concepts.
Practice and training	Repeated training and practice which provided opportunities for activating and applying metacognitive knowledge and metacognitive skills in multiple tasks, problems and contexts.
Teacher led metacognitive discussions	Whole class and group discussions of thinking and learning processes, discussions in which the teacher talks with their students about their thinking and learning in order to encourage and develop metacognitive thinking.
Student led metacognitive discussions	Student-led discussions were scaffolded using various instructional aids such as cues and prompts that evoke cognitive or metacognitive processes, scenarios and case studies, stimulated recall using think aloud protocol, student teacher metacognitive modelling.
Explicit metacognitive instruction	Explanations and demonstrations by the teacher regarding specific cognitive or metacognitive strategies and knowledge construction activities which require learners to construct their metacognitive knowledge. Explicit instruction was also embedded in the student metacognitive journal and homework activities.

In this study quantitative and qualitative methods were combined to explore the classroom's learning environment, and the participants' metacognitive regulation. Previous studies have argued for the value of employing a mixed methods methodology in metacognitive research and many studies in the field of learning environments and across education employ such approaches (Johnson & Onwuegbuzie, 2004; Tashakkori & Teddlie, 2010).

Qualitative methods

The visual and audio recordings and student generated work were analysed to identify the strategies used, the frequency of such strategies and to relate these with the quantitative results obtained. The audio and video recordings were analysed to identify specific instances where events occurred during the class that specifically highlighted metacognitive regulation strategies. A rubric was developed for analysing visual and audio recordings for the occurrence of teacher actions (e.g., teacher dialogue, specific teaching and learning strategies) which aimed to promote metacognitive regulation in students (Table 3). Each of the teacher actions were further categorised as specifically aimed at development of student sub-categories of metacognition i.e., Planning, Monitoring, and Evaluation. Development of the rubric for analysis of visual and audio recordings included cross coder agreement involving the researcher/practitioner and the supervisors of this research project, who are professors in STEM education. The qualitative data collected was analysed using QSR International's NVivo 12 qualitative data analytical software.

Table 3. Summary of the rubric developed for video analysis teacher actions incorporate metacognitive regulation

Subcategory	Description of Behaviour
Planning	Teacher decides (or asks students to decide) on procedures necessary for performing the task, individually or with others Teacher sets or clarifies (or asks students to set or clarify) task demands and expectations Teacher (allocates or asks students to allocate) individual roles and negotiates responsibilities Teacher sets (or asks students to set) goals and targets Teacher decides (or asks students to decide) on ways of proceeding with the task Teacher seeks (or asks students to seek) and collect necessary resources
Monitoring	Teacher assesses (or encourages students to assess) the quality of task performance (of self or others) and the degree to which performance is progressing towards a desired goal Teacher (or teacher asks students) to review progress on task (keeping track of procedures currently being undertaken and those that have been done so far) Teacher rates (or teacher asks students to rate) effort on-task or rates actual performance Teacher rates (or teacher asks students to rate) makes comments on currently memory retrieval Teacher checks (or encourages students to check) their behaviours or performance, including detection of errors, self-corrects.
Evaluation	Teacher verbalises or behaves, (or asks a student to verbalise or behave) to reviewing task performance and evaluating the quality of performance (by self or others) Teacher explains (or asks students to explain) their own learning or explain the task. Teacher evaluates (or asks students to evaluate) the strategies they used Teacher rates (or asks students to rate) the quality of their performance Teacher observes or comments (asks students to observe or comment) on task progress, or to test the outcome or effectiveness of a strategy in achieving a goal

Quantitative methods

The Metacognitive Awareness Inventory (MAI) is a self-report instrument designed to assess general self-regulated learning skills across the disciplines. Developed by Schraw and Dennison (1994), the MAI was developed specifically to address the two theoretical components (or dimensions) of metacognition: knowledge of cognition (17 items) and regulation of cognition (35 items). The analysis quantitative data collected in this included the use of the statistical software Statistical Product and Service Solutions (SPSS) (IBM Corp., 2017).

RESULTS

Visual and audio recording Analysis

Using the rubric developed the number of occurrences of teaching and learning strategies which address metacognitive regulation are displayed in Figure 1.

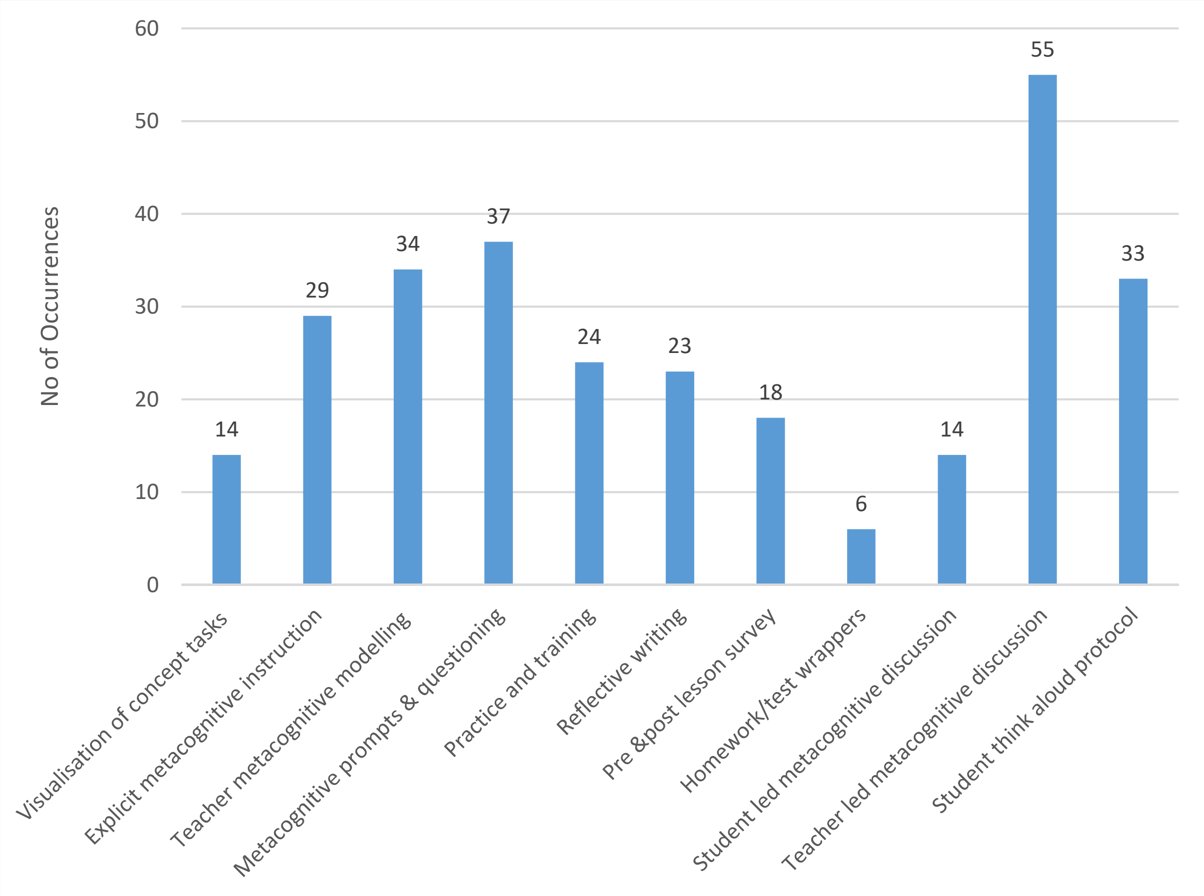


Figure 1. Occurrences of teacher actions addressing metacognitive regulation

Using the rubric developed in this research, the number of occurrences (during the 40 lessons analysed) of indicators of actions or dialogue by the teacher which addressed the subcategories of metacognitive regulation are outlined in Figure 2.

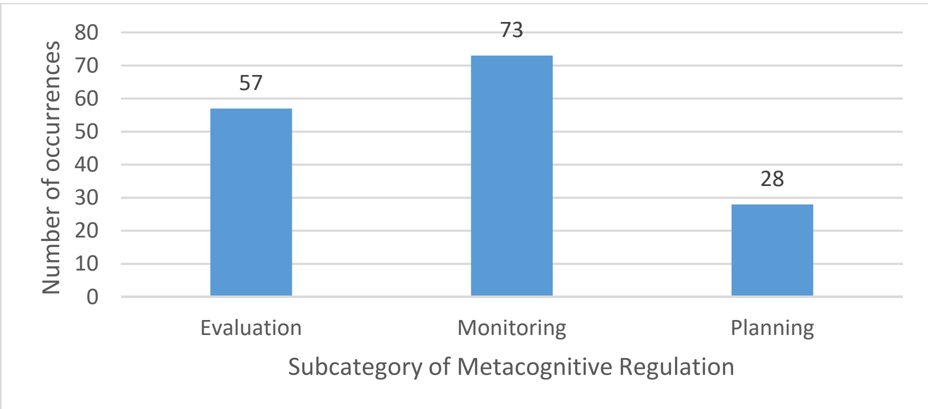


Figure 2. Occurrences of teacher actions addressing metacognitive regulation

A paired sample t test was conducted to evaluate whether a statistically significant difference existed between the mean scores in the subcategories of metacognitive regulation in the Metacognitive Awareness Inventory (MAI) before and after of the year. For the purpose of this study, only the 35 items which pertain to metacognitive regulation are included. A summary of the statistically significant increases in the subcategories of metacognitive regulation (MAI) scores are shown in Figure 3. The subcategory of metacognitive regulation (using the MAI) which did not show a statistically significant change in score pre and post the data collection period was evaluation.

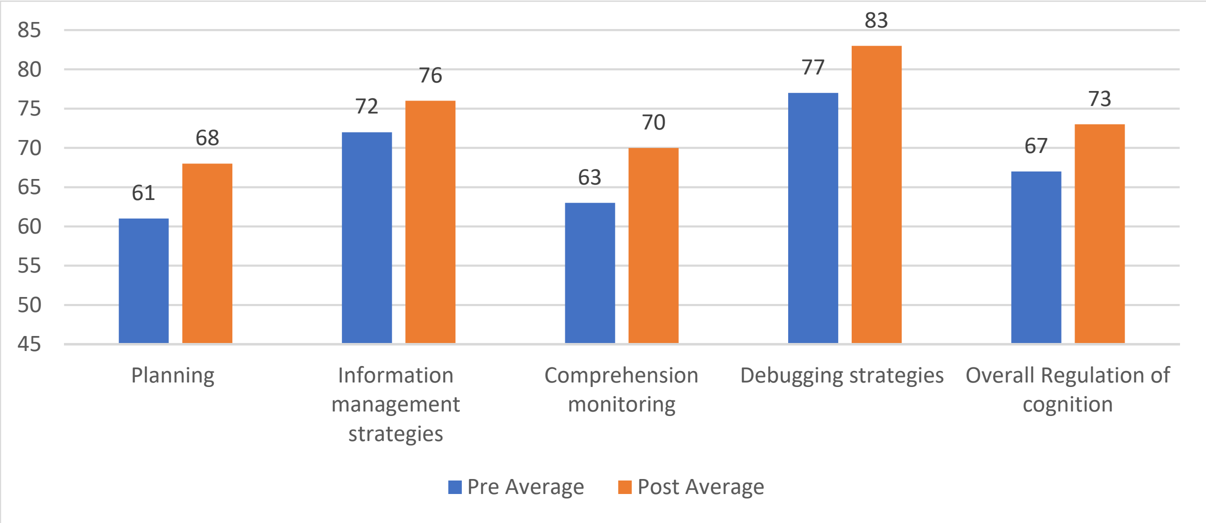


Figure 3. Metacognitive regulation subcategories within Metacognitive Awareness Inventory (MAI) which showed statistically significant changes pre and post the research period.

The paired sample t test which was conducted to evaluate whether a statistically significant difference existed between the mean metacognitive regulation (MAI) scores before and after the research period showed there was a statistically significant difference. Assumption testing indicated no gross violation of the assumptions. The results of the sample t test where significant $t(23) = -4.15$, $p < 0.05$, $\eta 2 = 0.43$ indicating that there is significant increase in metacognitive regulation (MAI) from pre survey (Mean = 0.7, S.D.=0.1, N = 24) to the post survey (mean = 0.7, S.D. = 0.07, N = 24). The effect size (= 0.85) was large based on Cohens conventions (1988). The mean increase was 0.065888 with the 95% confidence interval for the difference between the means of 0.1 and 0.03. The researcher rejected the null hypothesis.

DISCUSSION AND CONCLUSIONS

The metacognitive teaching plan implementation was analysed using the visual and audio records of 40 lessons throughout the academic year 2019-2020. The developed metacognitive journal was used in lessons throughout the research period. Development of the rubric for the analysis of the visual and audio recordings required extensive time-consuming work to ensure validity and reliability using cross coder agreement.

Analysis of the pre and post research period surveys showed a significant statistically important increase in student metacognitive regulation at the beginning and at the end of the research period, as measured by the MAI. The metacognitive teaching plan which included teaching and learning strategies to promote metacognitive regulation identified in the literature were embedded in lessons in upper second level chemistry lessons, this was evaluated and confirmed using visual and audio analysis.

The metacognitive teaching and learning strategies which address metacognitive regulation used in this study were evaluated using visual and audio recordings, practitioner reflections and pre and post research period surveys. A rubric was developed for analysis of visual and audio recordings of teacher actions which incorporate metacognitive regulation

There was a statistically significantly important increase in student's scores in the overall knowledge of cognition, regulation of cognition as reported using the Metacognitive Awareness inventory (MAI).

Limitations of this research include there are that additional factors which could have had an influence of the average metacognitive regulation score (as assessed by the MAI) of the students in the research include the metacognitive teaching plan, increases in conceptual understanding and confidence in chemistry and increase in the age and experience of the students. Isolating any of these variables in a quasi-experimental investigation would be difficult to achieve, and have ethical implications if such a study was carried out.

Future work

This study is part of a larger research project. The research period for the second phase in this research took place during the academic year 2020-21 which involved a pre-service chemistry teacher training module. Phase two of the project was informed by the research carried out in current study. Phase two of the research focused on two aspects of preservice education, namely (i) developing the preservice teachers own metacognitive knowledge and regulation, and (ii) developing the preservice teachers' metacognitive teaching and learning strategies which they can use to develop greater conceptual understanding of chemistry in their future students. A full analysis of the data collected in phase two of the research is currently being carried out.

REFERENCES

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of educational research*, *65*(3), 245-281.

Chiu, J. L., & Linn, M. C. (2012). The role of self-monitoring in learning chemistry with dynamic visualizations. *Contemporary Trends and Issues in Science Education, 40,* 133-163.

Cleary, T. J. (2011). Emergence of Self-Regulated Learning Microanalysis: Historical Overview, Essential Features, and Implications for Research and Practice University of Wisconsin–Milwaukee. In *Handbook of self-regulation of learning and performance* (pp. 343-359). New York, Routledge.

Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. Metacognition: Knowing about knowing, 207, 226.

Elder, L., & Paul, R. (May 01, 1998). The Role of Socratic Questioning in Thinking, Teaching, and Learning. *The Clearing House: a Journal of Educational Strategies, Issues and Ideas, 71,* 5, 297-301.

Ellis, A. K., & Denton, D. (2013). *Teaching, learning, and assessment together: Reflective assessments for middle and high school mathematics and science*. Larchmont, NY, Routledge.

Flavell, J. H., Miller, P. H., & Miller, S. A. 2002. Cognitive development (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Georghiades, P. (2004). From the general to the situated: Three decades of metacognition. *International Journal of Science Education*, 26(3), 365-383.

Gilbert, J. K. (2005). Visualization: A metacognitive skill in science and science education. In *Visualization in science education* (pp. 9-27). Dordrecht, Springer.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist*, 53(4), 449.

Hofstein, A., Kipnis, M., & Kind, P. (2008). Enhancing Students' meta-cognition and Argumentation Skills. In *Science education issues and developments*, (pp.*59-94*). New York: Nova Science Publishers.

IBM Corp. (2017). *IBM SPSS Statistics for Windows*. Armonk, NY: IBM Corp. Retrieved from https://hadoop.apache.org

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *Journal of Educational Psychology, 83,* 3, 1991.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Review*, *10*(1), 57-74.

Muis, K. R., Winne, P. H., & Jamieson-Noel, D. (2007). Using a multitrait-multimethod analysis to examine conceptual similarities of three self-regulated learning inventories. *British Journal of Educational Psychology*, *77*(1), 177-195.

Pieschl, S., Stahl, E., Murray, T., & Bromme, R. (2012). Is adaptation to task complexity really beneficial for performance?. *Learning and Instruction*, *22*(4), 281-289.

Pintrich, P. R., Wolters, C. A., & Baxter, G. P. (2000). Assessing metacognition and self-regulated learning. in *Issues In The Measurement Of Metacognition*, ed. Gregory Schraw & James C. Impara. Lincoln, NE: Buros Institute of Mental Measurements.

QSR International Pty Ltd. (2018) NVivo(Version12), https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Rickey, D., & Stacy, A. M. (2000). The role of metacognition in learning chemistry. *Journal of chemical education*, *77*(7), 915.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco, Jossey-Bass.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional science*, *26*(1), 113-125.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460-475.

Stensvold, M., & Wilson, J. T. (1992). Using concept maps as a tool to apply chemistry concepts to laboratory activities. *Journal of Chemical Education*, *69*(3), 230.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social & behavioral research.* Los Angeles, Sage.

Thomas, G. P. (2012). Metacognition in science education: Past, present and future considerations. In (1st Ed), *Second international handbook of science education* (pp. 131-144). Dordrecht, Springer.

Thomas, G. P., & McRobbie, C. J. (2001). Using a metaphor for learning to improve students' metacognition in the chemistry classroom. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *38*(2), 222-259.

Tien, L. T., Rickey, D., & Stacy, A. M. (1999). The MORE thinking frame: Guiding students' thinking in the laboratory. *Journal of College Science Teaching*, *28*(5), 318.

Trimble, J., Lovatt, J., & Finlayson, O.E. (2019). Questioning strategies to promote and assess, cognition metacognition and conceptual understanding in chemistry: an analysis. In Levrini, O. & Tasquier, G. (Eds.), Electronic Proceedings of the ESERA 2019 Conference. The beauty and pleasure of understanding: engaging with contemporary challenges through science education, Part 1 (co-ed. Anna De Ambrosis & Odilla Finlayson), (pp.147 - 153). Bologna: ALMA MATER STUDIORUM – University of Bologna. 978-88-945874-0-1978-88-945874-0-1

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., ... & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and learning*, *4*(1), 63-85.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of self-regulation* (pp. 531-566). San Diego, Academic Press.

Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. *Studies in Science education, 49*(2), 121-169.

THE IMPACT OF PHYSICS ANXIETY ON PERSONAL AND COLLECTIVE AGENCY IN THE PHYSICS CLASSROOM

John Connolly

UCL, Institute of Education, London, UK

School students' attitudes towards physics have been well documented and it is commonly perceived that female students' attitudes towards physics are less positive than those of male students. This has been found in studies researching a range of factors including interest, identity, subject choice, self-beliefs and participation. Physics anxiety is a phenomenon that can be described as students having negative emotional dispositions towards activities in physics lessons, and physics lessons themselves beyond test anxiety. Using Bandura's social cognitive theory, one of the antecedents of self-efficacy, emotional arousal, which relates to levels of anxiety, significant statistical differences were found between the anxiety levels of female students and their male peers in physics lessons. Observations of physics lessons and subsequent qualitative interviews with students confirmed that female students tended to minimise their participation in physics lessons due to concerns about obstructing the learning of others if they proffered incorrect answers in lessons or did not understand a concept. High attaining female students highlighted that physics was not a subject for them confirming widely held stereotype that physics is viewed as more of a male pursuit.

Keywords: Physics, Self-Efficacy, Gender Issues

BACKGROUND

With fewer students choosing physics compared to the other sciences for post compulsory education and the generally perceived unpopularity of physics as a subject or topic of interest in wider society, there is a need to discover the essence of students' relationship with physics in school. Female students' attitudes towards science in school are typically less favourable than male students' attitudes (Barmby, Kind and Jones, 2008). Mathematics anxiety is a term that describes how students hold negative affective states/expressions in anticipation of and during lesson activities in mathematics (Roth and Walshaw, 2015). Roth and Walshaw (ibid) used Vygotsky's socio-cultural framework to investigate maths anxiety in the classroom and highlighted how students' anxiety stems from a loss of agency within the context of the classroom due to the influences of the classroom environment and the relationships students have with their teacher and peers. Though less researched as a phenomenon, physics anxiety has been found to similarly impact on students' affective states in physics lessons and has been considered as significant as other factors such as students' self-efficacy (Sahin *et al*., 2015). Student anxiety in physics has been labelled as a 'stereotype threat' (Steele and Aronson, 1995), highlighting students' perceptions of physics being a difficult subject as well as a male dominated pursuit. Trujillo and Tanner (2014) suggest moving from a student deficit model that solely focuses on conceptual development to a dynamic model that takes into account affirmative affective influences in order to improve students' experiences of their learning environment that will also help to develop their conceptual understanding. The importance of emotions in learning and development was acknowledged by Vygotsky when considering the collaboration involved in the interpsychological relationships as part of the Zone of Proximal Development (Mahn and John-Steiner, 2002). Positive emotions help to develop confidence that can affirm one's beliefs about their capabilities that enhances one's agency that can influence goal driven actions.

Bandura's (1986) Social Cognitive Theory (SCT) attempts to explain the influencing factors that underpin human agency and its role in people's ability to exercise control in their lives. He argued that human agency derives from a dynamic relationship between personal, behavioural and environmental determinants. In the classroom, students' experiences of their surroundings and their relationship with their teacher and peers interact with personal factors that influences their actions in the classroom. Self-efficacy, a key personal determinant, is described as a person's situated belief that they can be successful in a specific pursuit. Bandura (1977, p. 3) defined self-efficacy as ". . . *beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments"*, which, in academic settings, relates to students' beliefs in their capabilities to succeed in subject related tasks and activities. A comparative study of male and female participants with similar performative levels in an introductory physics course (Marshman *et al*., 2018) reported that female students had lower self-efficacy than their male counterparts due to their experiences, and perceptions, of their course. The authors highlight structural and sociocultural biases as well as the impact of the physics class environment with raised anxiety levels leading to female students doubting their capabilities to be successful in physics.

Students' low self-efficacy beliefs can raise their anxiety levels whilst attempting challenging activities which can generate avoidance strategies (Britner and Pajares, 2006). Studies at high school level have shown gender differences in self-perceptions of capabilities in physics (Hazari *et al*., 2010). Hazari *et al.* (ibid) also identified student participation in physics lessons as a factor in enhancing students' self-perceptions, in part due to the sense of becoming an authoritative participant in class with greater confidence and empowerment, i.e. having a greater sense of personal agency.

Collective agency is another factor of human agency in Bandura's SCT and, in a school context, is associated with the interrelated dynamics of the classroom and the wider structures of school and society. In a physics lesson, the collective agency of the class, i.e. the teacher, students and support staff, is not simply the additive sum of personal agencies but an emergent group-level collective property (Bandura, 2001). Students' motivations and actions will be facilitated by the collective belief of the class that it is possible to understand physics concepts, be focussed in, and enjoy physics lessons. Bandura warns against a dualistic dichotomy between personal and collective agency. Humans are not abstract entities divorced from social systems. Individuals' interpretation of socio-structural authorised rules, practices and sanctions allows for a dynamic interplay between an individual and their environment. However, individuals do not just react to their environment, they operate proactively and generatively to influence and shape social structures highlighting the reciprocal nature of intrapersonal and interpersonal domains. Thus, if students, in particular female students, perceive their self-efficacy and personal agency as low in a physics lesson, the overall collective agency of the class will be impacted upon and may result in an overall reduced sense of capability to succeed in physics in school. This study's aims were to investigate female students' anxiety levels and their behaviours in physics lessons demonstrating their personal agencies in the wider social context of the class as part of a wider study focussed on interventions to enhance students' self-efficacy beliefs in physics lessons.

METHODS

A mixed methods approach was used in this study involving three methods: a survey, lesson observations and individual semi-structured interviews. 117 students (female 54, male 63) across five classes (aged 14-16 years old) of the same secondary school in London completed a self-efficacy survey during one of their physics lessons and one class of Year 10 students (15 years old) was chosen to observe physics lessons after the students had completed the survey. Semi-structured interviews were conducted with a sample of students from the Year 10 class after the lesson observations.

Self-Efficacy Survey

The survey instrument used was derived from the Sources of Self-Efficacy in Science - Physics (SOSESC-P) survey (Fencl & Scheel, 2005). The survey contains 33 items in total. The items are specifically designed to be grouped into four subscales for the theoretically derived four sources of self-efficacy (Bandura, 1997) (10 items for Mastery Experiences, 7 items for Vicarious Learning, 7 Items for Social Persuasion and 9 items for Emotional Arousal). This study concentrates on the nine items explicitly focussed on the emotional arousal antecedent of self-efficacy. The surveys were conducted in physics lessons and participants responded to each item using a 10-point Likert scale with responses from 1 (not at all like me) to 10 (very much like me). Some item's scores were reversed. For example, for item 18 'Physics makes me feel uneasy and confused', a score of '2' indicating that the student did not think this was like their behaviour was reversed to a score of '9'. The average score of all the 33 items was calculated to provide a proxy self-efficacy score in physics in school out of 10, the higher the score the higher the level of self-efficacy. Average scores for each antecedent were also calculated, this research utilised the data for the emotional arousal antecedent. Descriptive and statistical analyses of the survey data were conducted on the whole scale and each of the antecedent subscales to explore if there were any significant differences between male and female students' responses. The internal consistency of the entire survey and each of the four antecedents was found using Cronbach's alpha and was found to be 0.912 for the entire survey and 0.820 for the emotional arousal subscale.

Lesson Observations

A high achieving Year 10 class (24 students, 9 female and 15 male) was selected for lesson observations. The class was taught by a male teacher with two years' experience working as a physics specialist at the school. After analysis of the survey data and a discussion with the teacher about seating arrangements in the class, a group of students (3 female and 3 male) were selected to be the focus of the observations. The author observed four physics lessons each of one hour duration. The Year 10 class were learning about different aspects of Forces for each lesson following the AQA awarding body specification.

The focus of the observations was to monitor the students' interactions with peers and with the teacher as well as to record the students' behaviours during the lesson activities. Fieldnotes were taken during each lesson with details of the episodes of the lesson recorded as well as details of the student interactions and behaviours. Immediately after each lesson, the author annotated the fieldnotes with reflections about the student interactions and engagement during the lessons. Codes were identified according to the self-efficacy antecedents of Bandura's (1997) SCT for the group as a whole and for the male and the female students respectively.

Interviews

Individual semi-structured interviews were conducted with the six students who had been the focus of the lesson observations. The interviews provided an opportunity for the students to elaborate on their responses to the survey as well as to discuss their experiences (including affective states) during the lessons that were observed. All interviews were audio recorded and subsequently transcribed by the author.

The author adhered to a constant comparative analysis (Miles, Huberman and Saldaña, 2020) between the data sets. The interview transcripts were carefully read and coded based on Bandura's (1997) construct of agency involving personal, behavioural and environmental factors. The codes for the interviews and lesson observations were compared to discover emerging themes about the students' affective experiences and the exercise of their agency in the lessons.

FINDINGS

Independent t-tests were conducted for the data for all 117 participants who completed the SOSESC-P survey comparing the scores for male and female students. There were significant statistical differences for 18 of the 33 items from the survey including 7 of the 9 items associated with the emotional arousal antecedent of self-efficacy (see Table 1). The majority of male and female students disagreed with the statement that it was fun to go into physics lessons but there was an overall positive response to conducting practical work in lessons. This is consistent with research on students' attitudes of science education (Barmby, Kind and Jones, 2008) and the affective value of practical work (Sharpe and Abrahams, 2020). Data from the lesson observations echoed the more positive attitude towards conducting practical work and watching demonstrations compared to working on questions in worksheets and exam style questions.

The data from the emotional arousal subscale of the survey highlight the overall more negative affective dispositions of the female students in physics lessons compared to the male students. Table 1 shows that female students have lower scores on average for each of the items for the emotional arousal antecedent of self-efficacy. The average scores for the emotional arousal antecedent of self-efficacy for the class was 6.24 with male students' scores higher than that of the female students (male - 6.50, female - 5.80). Negative responses for the majority of students were particularly noted in relation to taking tests and when solving problems in class. The female students' response to worrying about solving physics problems was statistically significantly different to the male students' response for the same item.

Table 1. Average scores of emotional arousal survey items.

Emotional Arousal Survey Items Sample	Male	Female
My mind goes blank and I am unable to think clearly during physics tests*	6.93	5.76
I usually feel at ease in physics lessons*	5.94	4.35
I enjoy physics experiments*	7.89	6.59
I usually don't worry or stress about solving physics problems**	5.19	4.22
Physics makes me feel uneasy and confused (R)*	6.75	5.27
I get a sinking feeling when I think of trying to answer difficult physics questions and problems (R)*	6.38	5.04
I can relax and enjoy physics lessons*	5.63	4.41
I think it is fun to go into physics lessons	4.71	3.69
I get uptight when I take physics tests	5.93	5.19

*t<0.01 **t<0.05, R – *Reversed item – scores for items marked (R) have been reversed*

Data from the lesson observations highlighted that generally female students, though on-task in all activities involving individual and small group tasks and discussions, did not engage to the same level as the male students. This was also witnessed in whole class discussions where the female students remained quiet and rarely voluntarily participated called upon to do so by the teacher. Female students, tended not to volunteer to answer the teacher's questions in front of the whole class by raising their hand but would answer questions from the teacher if directly called upon to do so. Conversely, the male students, engaged very willingly, particularly to answer the teachers' questions. The male students did not appear to be deterred by providing an incorrect response and remained keen to respond to subsequent questions.

There was a difference in how male and female students overcame challenges during activities. If a male student could not answer a question on a worksheet, they would actively seek to find a correct method or solution by asking a peer. On the other hand, female students tended to either move on to another question or disengage from working on the worksheet rather than ask a peer for support. If the teacher noticed that the female student had stopped working, he would ask if they needed support and the student would ask for help.

During practical activities, if working in a mixed groups the male students would take charge of organising the equipment. One activity involved measuring the acceleration due to gravity. The female students stood back when the male students assembled the equipment but did get involved in discussions about recording appropriate measurements and finding patterns in the data as the practical was being conducted. When the group was asked to report their data to the whole class, one of the male students opted to state the group's results with the female students not willing to volunteer.

As well as answering the majority of questions during episodes of whole class discussion, male students asked more questions directed towards the teacher. This resulted in explanations from the teacher and other male students in response to the questions but the female students did not partake in these discussions. These episodes were quite lively and the male students appeared to enjoy the back and forth discussion with the teacher. The female students did appear to be attentive to these discussions, as confirmed in the interviews, but were not willing to interject or ask for further explanation if they did not understand an aspect of the discussion relating to content on forces.

The lesson observations highlighted that the students were having very different experiences during physics lessons and their agency was apparent whether they were visibly fully engaged or appearing to be passive during the lessons. Although the teacher had planned for many collaborative activities during the lessons with space for input from the students, female students worked with their peers less compared to the male students. The male students appeared to be happy to work collectively whilst the female students, working together during practical activities, mostly worked independently for other activities.

During the interviews, when asked about their lack of participation during whole class lesson episodes, the female students' responses included that they did not want to obstruct their peers' learning and did not want to answer incorrectly as they were worried about how this would be viewed by their peers.

I think I am more keen to do practicals rather than joining in in class discussions because I feel a lot of pressure, like getting answers wrong and not understanding it as well as other people in the classroom (not understanding).

They made comparisons with their peers which led to judgements about their own capabilities to engage in the physics activities and their ability to be successful in the subject at school.

I just feel like, I don't know what the word is but my self-esteem would go low. It's like everyone else understands it (Physics) *and I don't understand it and I am holding up the class because I don't know the answer.*

The female students expressed that they had high anxiety levels in school on the days that they had physics lessons and had thus developed a 'coping mechanism' of low participation levels in class in an attempt to minimise their feelings of anxiety. Two female students stated that they were self-motivated to learn physics at home by researching for appropriate YouTube videos on the physics topics that were taught in lessons. One female student, although able to achieve well in physics tests due to time spent revising at home, did not expect to be successful in physics overall, and declared that if she had an option of no taking the subject, she would chose an alternative such as chemistry.

In contrast, the male students expressed satisfaction with their engagement and understanding of concepts in the forces topic during the physics lessons and their ability to support others.

I do well in physics when I fully understand the concept and I have realised that I can explain that concept to anyone in my whole class and I understand it well. That's when I think okay, I have learnt that well, I'm good at it.

They also discussed that they were confident in their capabilities to be successful in the subject and were not put off if they made a mistake and recalled what they had done to ensure they remembered the correct answer in the future by collectively working with one another student,

I can explain to Ajay (pseudonym) *and let's say I got something wrong that Ajay knows, he will correct me. The concept that I know thoroughly, that I think I know thoroughly, but let's say I missed out a very important part of it, he'll explain it back to me as well, so we both are like, our minds are coming together.*

The same student had provided a wrong answer for a question from the teacher during a summing up of a worksheet activity. He appeared disappointed that he had gotten the question wrong but, when appropriate, he sought an explanation from a peer and appeared to be satisfied that he now understood his mistake.

The male students explained that they put more effort to learning physics in school with only a simple going over of content at home to ensure that they understood the concepts taught that day. As one student stated,

To rate out of ten, I probably give a seven in class and then out of school probably four… I focus on it in class and go over it once at home.

The male students demonstrated that they had good working relationships with one another with an apparent collective agency that aided their support for one another to learn physics in lessons. The support did extend to the female students to but requests for this support were not forthcoming from the female students. Their beliefs about their capabilities and their agency meant that they had frequent input into the lesson and spent more time interacting with peers and with the teacher than the female students did.

CONCLUSION

This research describes how students exercise their agency in lessons based on their beliefs about their capabilities in relation to the context of the class and the activities they are asked to do in order to learn physics. The findings of the analysis of the survey data echo the findings of Sahin *et al*. (2015) in their evaluation of the Physics Anxiety Rating Scale where female students were found to have higher physics related anxiety than their male counterparts. Interviews with the female students demonstrated that although their anxiety about physics before and during lessons resulted in the female students working more independently, or perhaps to put it more aptly, in isolation, the students' lack of control of the conditions within the lessons (their agency) limited their actions during the lessons (Roth and Wilshaw, 2015). However, for some female students, this meant that they were resolved to supplement their learning by independently using resources from the internet at home.

Though their self-regulated skills appear to be nurtured by this working at home, female students are missing out on opportunities to develop a relationship with their physics teacher by not receiving as much direct feedback about their progress compared to their male peers who are keen to interact with the teacher as shown by the lesson observations. It could be argued that this lack of participation by capable students renders interactive teaching methods into a more transmission-based style of teaching for the student. With the students feeling anxious about participating and thus developing feelings of isolation during the lessons they have reduced opportunities to develop their sense of agency and take greater control of their learning during lessons. The lack of occasions to utilise their interpersonal skills in physics lessons with their peers and teachers leads to doubts about their capabilities to succeed in physics. This attitude, that physics is not for them, reinforces the stereotype threat where they view the subject as difficult and a pursuit for others. This reduces the sense of collective agency in the class as well as the belief that everyone can contribute to the success of the class in learning physics.

There appeared to be two simultaneous lessons happening within the class. One, shaped by the actions of the male students which had attributes of an interactive lesson with engagement from the male students and the teacher with a strong sense of individual and collective agency. Whilst the other lesson was of more isolated working with little interdependence and low levels of individual agency and collective agency between the female students and the teacher. The female students' 'lesson' appeared to be derived from the male students' 'lesson', that is, their perceived lack of participation was determined by their experiences within the context of a more dominant 'lesson' of confident male students engaging and interacting with one another.

To support all students so that they have more positive perceptions of their capabilities and to develop more inclusive environments to enhance students' sense of personal and collective agency in physics lesson, their affective states in lessons need to be taken into consideration. The class cultural climate in physics lessons needs to provide space for students with low agential and capability beliefs to be able to engage in an environment where they feel that they can actively participate in lessons and reduce their isolation to develop a more supportive and collective approach to learning physics. A heightened sense of agency and feeling of belonging may engender an environment where more students consider that they have control when learning physics and that physics might be a pursuit that they are able to successfully participate in.

REFERENCES

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2).

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Bandura, A. (2001). Social cognitive theory: an agentic perspective. *Annual review of psychology*, *52*, 1–26.

Barmby, P., Kind, P. M. & Jones, K. (2008). Examining changing attitudes in secondary school science. *International Journal of Science Education*, *30*(8), 1075–1093.

Britner, S.L. & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485–499.

Fencl, H. & Scheel, K. (2005). Engaging students. *Journal of College Science Teaching*, *35*(1), 20.

Hazari, Z., Sonnert, G., Sadler, P. M. & Shanahan, M. C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching, 47*(8), 978-1003.

Mahn, H. & John-Steiner, V. (2002). *The Gift of Confidence: A Vygotskian View of Emotions. In Learning for Life in the 21st Century*. Oxford, UK: Blackwell Publishing Ltd, pp. 46–58.

Marshman, E., Kalender, Z., Nokes-Malach, T., Schunn, C. & Singh, C. (2018). Female Students with A's Have Similar Physics Self-Efficacy as Male Students with C's in Introductory Courses: A Cause for Alarm? *Physical Review Physics Education Research*, *14*(2).

Miles, M. B., Huberman, A. M. & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook*. Newbury Park, CA: Sage.

Roth, W. M. & Walshaw, M. (2015). Rethinking affect in education from a societal-historical perspective: The case of mathematics anxiety. *Mind, Culture, and Activity*, *22*(3), 217-232.

Sahin, M., Caliskan, S. & Dilek, U. (2015). Development and Validation of the Physics Anxiety Rating Scale. *International Journal of Environmental and Science Education*, *10*(2), 183-200.

Sharpe, R. & Abrahams, I. (2020). Secondary school students' attitudes to practical work in biology, chemistry and physics in England, *Research in Science & Technological Education, 38*(1), 84-104.

Steele, C. M. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of personality and social psychology*, *69*(5), 797.

Trujillo, G. & Tanner, K. (2014). Considering the Role of Affect in Learning: Monitoring Students' Self-Efficacy, Sense of Belonging, and Science Identity. *CBE - Life Sciences Education*, *13*(1), 6-15.

KNOWLEDGE OF PHYSICS AND SUCCESS IN COMMENTING ON SILENT VIDEO TASK

Simona Gorčáková[1], Klára Velmovská[1]

[1] Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic

According to the European Union, communication is one of the key competencies. (Arjomand et al., 2013) Therefore, it is important to develop communication skills during the education. The space for this development could be provided by silent video task. The paper contains a definition of the silent video task and a survey aimed at correlation between knowledge of physics and success in commenting on silent video, which was conducted among high school students.

Keywords: Silent Video, Verbalization, SOLO taxonomy

INTRODUCTION

One of the aims of science education in Slovakia is to develop students' communication skills, i.e., to develop the formulation and expression of their thoughts and opinions. (ŠPÚ, 2015, p. 7). Verbalization of thoughts is an everyday part of a life, not just school life. To verbalize means to express thoughts, opinions, or emotions through words. (Cambridge Dictionary, 2020) For a teacher who wants to develop students' cognitive processes, it is important to know their thoughts. Therefore, the student's ability to verbalize their thoughts is crucial.

The development of the ability to verbalize one's own ideas begins in kindergarten and continues. "By constantly guiding the child to formulate his or her own ideas, his or her ability to generalize is developed. Even though the child implicitly perceives that he understands the observed, it is not easy for him to translate his cognitive idea into a verbalized expression. Much information is brought into the mind without verbalization and completes the understanding of a certain phenomenon without speech coding. That is why expressing one's own ideas is quite demanding." (Žoldošová, 2013) Verbalization of thoughts is closely related to communication skills.

There are current trends in education which are aimed at prioritizing the development of competencies over the acquisition of knowledge. In 2006, the European Union adopted a document entitled Key Competences for Lifelong Learning - A European Framework of Reference (EU, 2006), which sets out eight key competences needed for personal satisfaction, active citizenship, social cohesion, and employability in the knowledge society:

1. communication in the mother tongue,
2. communication in foreign languages,
3. mathematical competence and basic competences in the field of science and technology,
4. digital competence,
5. learn to learn,
6. social and civic competences,

7. initiative and entrepreneurship,

8. cultural awareness and expression.

Communication skill is therefore an important competence that needs to be developed. "The information and communication aspect of scientific thinking is revealed in skills such as communicating and sharing knowledge, publishing discoveries, working cooperatively as a team, participating in congresses, discussing theories and solutions with colleagues, evaluating one's own and others' communications and assuming personal and social responsibilities." (Manassero-Mas, Vázquez–Alonso, 2020).

The opportunity of verbalizing the thoughts and developing the communication skills could be given by silent video task. Silent video is a short video, less than 2 minutes long, which shows a physics phenomenon or experiment. The task for student is to prepare and record the voice-over to the video. It turns out that silent video can contribute to the development of communication skills, argumentation, critical thinking, or evaluation skills. Silent video can be also used as an instrument of determining the communication skills level.

Since the aim of science education is also the development of communication skills and in order to develop communication skills, it is necessary to know their current level, we were interested in the level of communication skills of our students in relation to their knowledge, so we conducted a survey in which the silent video task was used.

The survey was conducted to find a correlation between communication skills and knowledge of physics and success in commenting on silent video. Therefore, it was necessary to evaluate the voice-overs on the silent video. A tool called SOLO taxonomy (Biggs and Collis, 1982) was used to evaluate them. The division of knowledge into five hierarchical levels is contained in SOLO taxonomy:

- *Prestructural level* – contains incorrect findings that are not related to the topic or that are irrelevant.
- *Unistructural level* – contains findings focused only on one phenomenon, one specific aspect described by the student.
- *Multistructural level* – although it consists of several findings, students only name them.
- *Relational level* – contains multistructural findings, which students relate to, combine, analyse, apply and create a broader view of the phenomenon.
- *Extended Abstract level* of findings contains a description of different new views on studied phenomenon, evaluation, generalisation and creation of new concepts.

Silent video task and SOLO taxonomy are important elements of conducted survey.

METHOD

The survey was focused on students' communication skills. The objective was to find out whether there is a correlation between knowledge of physics and success in commenting on silent video. Hypothesis was set as follows:

H: *The higher the level of student's knowledge of physics, the higher the score will be obtained in silent video task by the student.*

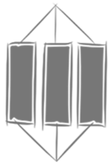
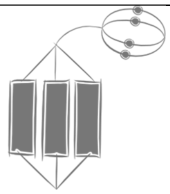
The survey was conducted with the students of 3rd year of French Bilingual High School in Bratislava. The number of participating students was 60. The task was to record the voice-over to silent video named *Ball*. The video shows a real situation approaching the behaviour of objects in the car. It is a video of a man throwing a ball in a car in three different situations. First, when the car does not move, then when the movement of a car is a uniform linear motion, and finally when the movement is non uniform linear motion with acceleration.

Figure 1. Video shot with QR code of the website containing the video – Ball.

The students' voice-overs were evaluated with the help of evaluation sheet (Fig. 2) using the SOLO taxonomy created by Biggs and Collis (1982). Since the SOLO taxonomy is a hierarchical division, we did not sum up points for individual levels in the sheet, but we chose one of the numerical options. For instance, if it was pronounced according to what they examined the movement, but the trajectory of ball was not described, total score was 2 points.

Table 1. Evaluation sheet with criteria created according to SOLO taxonomy.

	It is stated that this phenomenon is related to the speed of the car, which is first immobile, then is in a uniform motion and finally in an accelerated motion.	1	
	It is pronounced, according to what we examine the movement, trajectory of the ball is described.	2 3	

	The trajectory of the ball is related to the speed of the car.	4 5	
	The speed of the ball is considered relative to the car, as well as relative to the road, it is also explained why in the third case the ball moves differently.	6	

Before carrying out the silent video task, the test with physics tasks was used to find out the level of physics knowledge. The main topic of the test was Motion because it is the key knowledge for creating voice-over to this video. Test was composed of several types of exercises i.e. open-ended questions, multiple choice questions, quantitative and qualitative task. Then the score of the test was compared with the score obtained for the voice-over.

RESULTS

All results were transformed in the graph (Fig. 2) with *x*-axis showing score expressed as a percentage of the test and *y*-axis showing the score expressed also as a percentage the student with the given score from the test obtained for his voice-over.

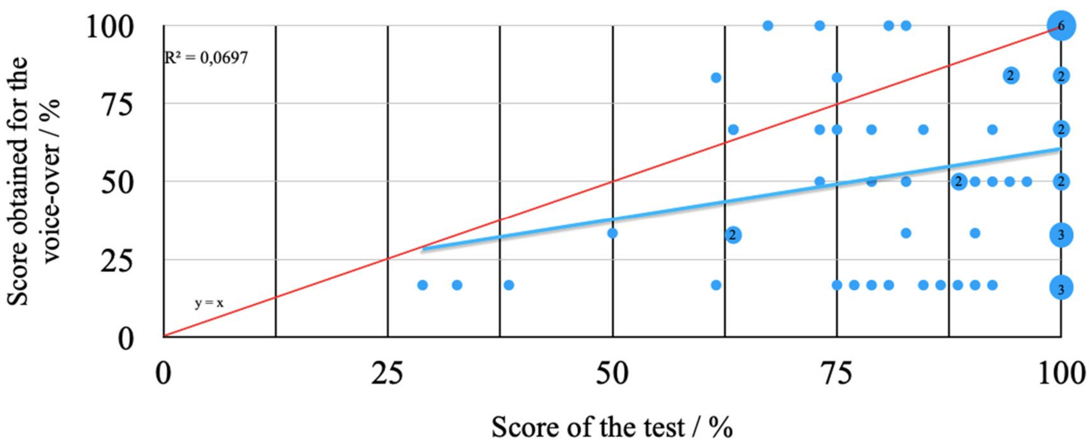


Figure 2. Graph of the correlation between physics knowledge and success in commenting on silent video. Larger points represent multiple students with the same score.

After counting the correlation coefficient (R = 0.26), it can be stated that with the selected students there is weak correlation between physics knowledge and success in commenting on silent video. The hypothesis H need to be rejected. The rejection of hypothesis led to setting hypothesis H': *The score of the test has no influence on the quality of the voice-over.* The test of contingency (Pearson's Chi-squared test with Yates' continuity correction) was used to verify the hypothesis H' and p-value = 0.25. Hypothesis H' cannot be rejected.

DISCUSSION AND CONCLUSIONS

It turned out that students whose score of the test were low, they did not obtain the high score for the voice-over. When creating a comment, students had to apply knowledge of the topic and since they do not have them (as demonstrated by the test), they could not be successful in creating a comment. As many as 46 students obtained less score for the voice-over than for the test. They are represented in graph (Fig. 2) by points, which are situated below the line $y=x$. It can be related to a non-standard type of task. However, up to 31 students whose score of the test was more than 63%, obtained for the voice-over the score with the decrease of 30% or more compared to their score of the test. We interpret this as showing that the students had sufficient knowledge of Movement but have not been successful in creating the voice-over. The results showed that there is a weak correlation between physics knowledge and success in commenting on the silent video. We think that it could be caused by insufficient ability of students to verbalize their thoughts. In order to develop students' ability to express their thoughts and communicate using the correct terminology, we must create the space in our physics lessons. It turns out that one of the suitable ways to provide such space is a silent video task.

ACKNOWLEGMENTS

This paper was elaborated with the support of the projects KEGA no. 013UK-4-2021 *Methodical materials focused on the systematic development of critical thinking* and UK/257/2021.

REFERENCES

Arjomand, G. et al. (2013). KeyCoNet 2013 Literature Review: Key Competence Development in School Education in Europe [online]. Available on: https://lnk.sk/gozh [10 January 2022]

Biggs, J. B., & Collis, K. F. (1982). Evaluation the Quality of Learning: The SOLO taxonomy (Structure of the Observed Learning Outcome). New York: Academic Press, 1982. 262 s. ISBN 0-12-097550-5.

Cambridge Dictionary [online]. Cambridge University Press, ©2020, Available on: https://lnk.sk/ir02 [10 January 2022]

EU (European Union). (2006). Recommendation of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning. [online]. In *Official Journal of the European Union*. L 394, 30.12.2006, p. 10 – 18. Available on: https://lnk.sk/hjh1. [10 January 2022]

Manassero-Mas, M.A., Vázquez–Alonso, Á. (2020). Science Education as Thinking: The Beauty of Scientific Thinking and Critical Thinking. In *Levrini, O. & Tasquier, G. (Eds.), Electronic Proceedings of the ESERA 2019 Conference. The beauty and pleasure of understanding: engaging with contemporary challenges through science education, Part [7/7] (co-ed. Maria Andrée & Kalypso Iordanou), (pp. 816 - 825). Bologna: ALMA MATER STUDIORUM – University of Bologna. 978- 88-945874-0-1978-88-945874-0-1*

ŠPÚ (Štátny pedagogický ústav). (2015). Štátny vzdelávací program pre gymnáziá. [online]. Bratislava: ŠPÚ, 2015. Available on: *https://lnk.sk/btus* [11 January 2022]

Žoldošová, K. (2013). Ciele a obsah primárneho prírodovedného vzdelávania. [online]. Trnava: Pedagogická fakulta Trnavskej univerzity. 2013. ISBN 978-80-8082-688-8. Available on: https://pdf.truni.sk/e-ucebnice/primarne-prirodovedne-vzdelavanie/data/5475e2c4-db99-47b1-84a9-5ae23b0a11c5.html?ownapi=1. [12 January 2022]

DOMAIN-SPECIFIC THEORIES OF INTELLIGENCE: HOW STUDENTS' MINDSETS IN PHYSICS CHANGE WITHOUT INTERVENTIONS

Laura Goldhorn[1], Thomas Wilhelm[1] and Verena Spatz[2]

[1]Goethe-University, Frankfurt, Germany

[2]Technical University of Darmstadt, Germany

Implicit theories about intelligence influence students' goal setting, their engagement in learning and their academic performance. Whether students believe in the fixedness versus malleability of intelligence, they are referred to as holding a fixed versus a growth mindset. Research shows the benefit of a growth mindset especially in an academic context, and how interventions can foster a growth mindset. But a surprisingly small number of studies examined the changeability of mindset over time without an intended intervention. This study aims to survey physics specific mindsets which students hold from the first year of physics lessons to graduation class. N=900 students from academic high schools in Germany participated in this survey and four different mindset types could be identified: a growth mindset, two types of fixed mindset (one holding the fixedness of general intelligence accountable for physics success and one believing in a special giftedness in physics) and a mixed mindset including those beliefs that do not fit the mindset theory. The change in the mindset distribution over the different grades indicate an influence of learning physics on the beliefs students hold about this subject. Also, as indicated in other mindset studies and research concerning different non-cognitive factors in physics, there is a gender-specific differentiation measurable in physics specific mindsets.

INTRODUCTION

Students vary in their beliefs about the nature of intelligence. Since these beliefs are often unconsciously held, they are also referred to as implicit theories of intelligence. Entity theorists believe that intelligence is a fixed trait, and one cannot change it (fixed mindset). Incremental theorists on the other hand, believe that intelligence is malleable and can always be developed through effort (growth mindset) (Dweck, 1999). Students with a growth mindset show a preference for challenging tasks; they choose learning goal tasks over performance goal tasks, and they show a mastery pattern of behavior and higher engagement in learning (e.g. Blackwell et al., 2007; Yeager et al., 2019). Research on mindset theory shows that one's beliefs about the fixedness or malleability of traits are not global. For example, Cheng & Hau (2003) report that students hold differentiated mindsets on intelligence, personality, creativity, and emotional intelligence. There are only a few studies examining students' mindsets in different academic domains, but they implicate a tendency of a more fixed view on STEM subjects (e. g. Cimpian et al., 2007; Jonsson et al., 2012; Leslie et al., 2015; Meyer et al., 2015). Also, some data reports gender-specific differences in implicit theories, e. g. Gunderson et al. (2013) found girls to be more likely to hold a fixed mindset and Archer et al. (2020) found that especially female students strongly believe in a natural ability in physics as a premise for success. While a majority of mindset research is focused on interventions to foster an incremental belief, there is only limited knowledge about changes in distributions of incremental and entity beliefs in different age groups, and especially domain specific. Since such knowledge would allow a more targeted setting of growth mindset interventions, the present study aims to investigate physics learners' mindset distributions in different age groups: How do students' domain specific beliefs about fixedness versus malleability of their physics-ability change without the influence of a mindset intervention?

THE STUDY

To address student's domain specific theories of intelligence in physics, N = 1606 students from different secondary schools in Hesse (Germany) participated in a pen and paper mindset survey between November 2019 and February 2020, while for the purpose of better comparability only the subgroup of N = 900 (430 of them female) students from the participating 12 academic high schools are considered. Students from seventh grade, which is the first year of physics lessons in Germany, to graduation class had been participating, to examine the whole age span of physics learners. The mindset survey contains the commonly used four items of the "Theories of Intelligence Scale" (Dweck, 1999), asking about students' beliefs about general intelligence, e. g. "You have a certain amount of intelligence and really can't do much to change it" (α = .80). To focus on the students' physics specific beliefs two more scales have been developed. The scale "Giftedness in Physics" consists of four items such as "You need a certain giftedness for being successful in physics" (α = .81). The other scale with seven items is labelled as "Effort in Physics" (α = .83) and contains statements like "Everyone can understand physics, you just have to put in enough effort" (Goldhorn et al., 2020).

To identify students' physics specific mindsets a hierarchical cluster analysis (Ward's method with squared Euclidean distances) was conducted using the average ratings of the three scales. Four types of physics specific mindsets can be identified, two of them are manifestations of a fixed mindset. Students with a *fixed mindset "general intelligence"* believe that general intelligence as a fixed trait is accountable for one's success in physics. Students with a *fixed mindset "giftedness in physics"* believe in a domain specific giftedness, a special talent in physics, that is necessary for one's success in this area. Students with a *growth mindset in physics* neither believe in intelligence as a fixed trait nor that one needs a special giftedness in physics to be successful. The fourth cluster is called a *mixed mindset* since it doesn't fit the characteristics of a growth or a fixed mindset.

RESULTS

Overall, 45.8 % of the participating students hold a growth mindset in physic, 16.2 % hold a fixed mindset "general intelligence" and 13.7 % a fixed mindset "giftedness in physics". 24.3 % of the students don't match the criteria for fixed or growth mindset and are therefore assigned to the mixed mindset. For more detailed results, table 1 shows the mindset distributions ordered by grade.

Table 6. Mindset distributions for different classes from seventh grade to graduation class.

	fixed mindset "general intelligence")	*fixed mindset "giftedness in physics"*	*mixed mindset*	*growth mindset*
7th grade	13.8 %	4.3 %	12.8 %	69.1 %
8th grade	16.2 %	13.4 %	26.9 %	43.5 %
9th grade	19.8 %	10.1 %	28.6 %	41.5 %
10th grade	17.9 %	19.0 %	24.4 %	38.7 %
introductory phase	9.7 %	16.7 %	13.9 %	59.7 %
graduation class	12.8 %	18.0 %	27.1 %	42.1 %

The youngest participants are in seventh grade, since physics lessons in Hesse (Germany) start in this grade. While 69.1 % of the seventh-grade students hold a growth mindset in physics after only a few months of learning experience in this subject, the growth mindset percentage decreases during the following years. The largest gap is between seventh and eighth grade (from 69.1 % to 43.5 %) and the minimum of the growth mindset percentage is in the last year of middle school (tenth grade) with only 38.7 %. During the same time, the number of students holding a physics specific fixed mindset (the so-called fixed mindset "giftedness in physics") is drastically increasing. Only 4.3 % of the participating seventh graders hold this mindset, but after just one year of physics learning the percentage of the fixed mindset "giftedness in physics" is up to 13.4 % in eighth grade. And while the growth mindset percentage has its minimum in tenth grade, the fixed mindset "giftedness in physics" is at its maximum with 19.0 % at the same time. Besides the changes in the mindset distribution during middle school also the percentages in introductory phase are noticeable. Introductory phase is the first year of the upper school level. The number of students holding a growth mindset increases in this grade to 59.7 %, the next highest percentage after starting physics classes. But this much stronger growth mindset does not seem to be long-lasting and in the next grade, graduation class, the growth mindset percentage is decreasing to 42.1 % which is not significantly higher than in middle school.

The following diagrams are showing the changing percentages of growth mindset (figure 1) and fixed mindset "giftedness in physics" (figure 2) comparing female and male students. The girls start physics lessons with a higher percentage of growth mindset (72.7 %) than the boys (63.2 %). In ninth grade the growth mindset percentage is the same for female and male students (42.1 %) and in higher classes there is a larger percentage of boys holding a growth mindset in physics.

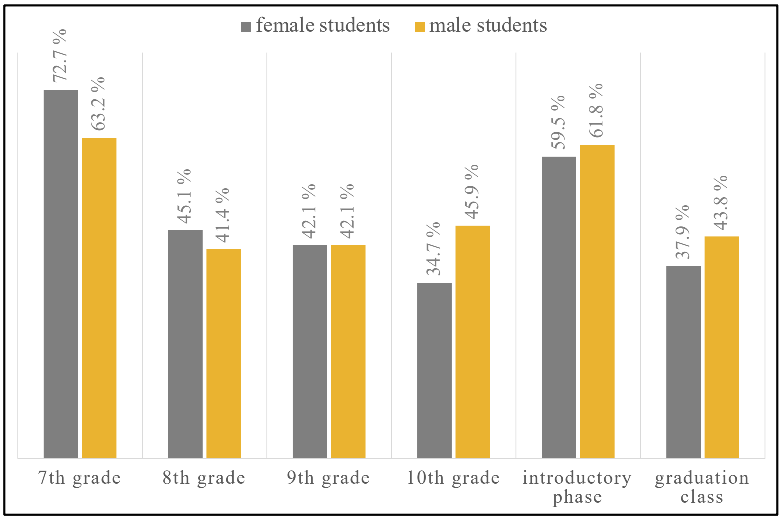


Figure 5. Growth mindset percentage for different grades, comparing male and female students.

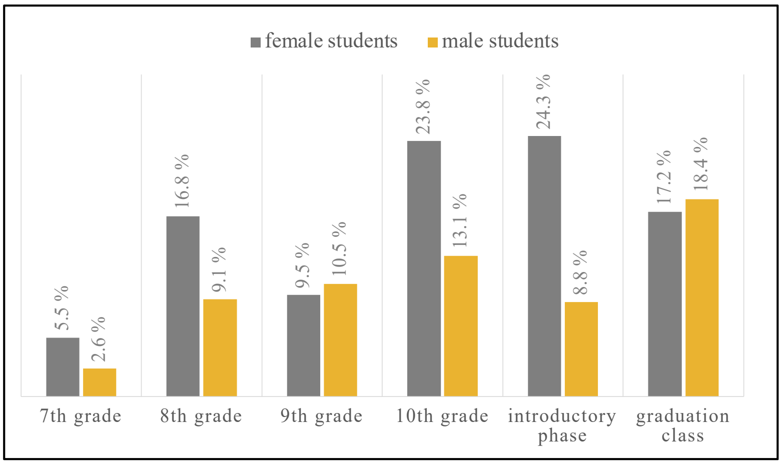


Figure 6. Fixed mindset "giftedness in physics" percentage for different grades, comparing male and female students.

DISCUSSION

Since there are students from seventh grade to graduation class participating in this study, the results show how students' mindsets in physics are changing over time. There are three main findings.

1. *Over the years the growth mindset percentage decreases and students show a higher percentage of fixed mindset in graduation class relatively to the first year of physics.*

While in seventh grade, when starting with physics lessons, there is a high percentage of students holding incremental beliefs (69.1 % in seventh grade), this growth mindset percentage decreases drastically over the next years. Similar findings are reported for domain-general mindsets, comparing the percentage of students with incremental beliefs in high school versus late elementary school (Cheng & Hau, 2003). Also, for domain-specific mindsets in math a decrease of the growth mindset is reported by Gunderson et al. (2017). Since there is some literature suggesting a higher percentage of fixed beliefs in STEM fields for adults (e.g. Jonsson et al., 2012; Meyer et al., 2015), monitoring mindset distributions over time is interesting in order to learn more about the development of the implicit theories. Our data indicates that this fixed mindset tendency in STEM is not given from the beginning but evolves with more exposure to physics lessons.

2. *The biggest mindset shift is observable during the first year of physics lessons.*

Looking more closely at the results of our study, the numbers show that there is a sharp decline of growth mindset percentage between the seventh and eighth grade (from 69.1 % to 43.5 %). Limeri et al. (2020) showed that undergraduate students' domain specific incremental beliefs decrease when faced with a challenging STEM-course. Physics lessons in general are described as challenging (e.g. Ornek et al., 2008), so middle scholars' beliefs may be influenced by starting physics classes in the same way. This, again would support the hypothesis that students' mindsets are influenced by physics lessons and are not just changing age-dependent during middle and high school.

3. *Girls hold stronger beliefs about an innate talent for physics.*

Comparing the gender specific mindset distributions, our results indicate that girls hold stronger beliefs about a fixed giftedness in physics (compare figure 2). And, as well as the sharp decline of the incremental belief during the first year of physics, there is a sharp increase in the domain-specific entity beliefs visible: the percentage of fixed mindset "giftedness in physics" rises from 5.5 % to 16.8 % for female students and from 2.6 % to 9.1 % for male students. These results are consistent with other studies asking about students' beliefs about talent in physics, e. g. Archer et al. (2020). In the first years of physics learning, the percentage of girls holding incremental beliefs is higher than the percentage of boys, but this changes from ninth grade onwards, when more boys tend to have a growth mindset. Several studies are supporting this result and found that boys held more incremental beliefs than girls (e.g. Gunderson et al., 2013). But there are also studies showing no gender specific differences (e.g. Gunderson et al., 2017). So, it will be important to examine this further and get more knowledge on the circumstances of possible gender differences in (domain-specific) mindsets.

CONCLUSION

In this study we reported students' domain-specific mindsets and how they change over time without any intervention taking place. Students' beliefs about the fixedness versus malleability of physics ability changes throughout their school years, and male and female students show a higher percentage of fixed beliefs in graduation class relatively to the first year of physics lessons. Even though there is no gender difference in the overall trend of an increasing domain-specific fixed mindset and a decrease of the growth mindset, girls seem to hold stronger beliefs about a fixed and innate talent in physics. The biggest shift in mindsets is visible between the seventh and eighth grade, while physics as a new school subject is introduced in seventh grade. Being faced with this new subject, often reported as challenging, seems to strengthen students' domain-specific entity beliefs about intelligence and ability. These findings are consistent with the results of other studies about fixed and growth mindset, especially with research concerning the STEM-field and mindset. Since the present study only collected data about physics specific mindset, we cannot say if the observable mindset changes over time are just age-dependent or being influenced by the academic setting and physics lessons in particular. Still, taken research about other non-cognitive factors into account (e. g. Hoffmann et al., 1998) our findings can be interpreted as another indication of physics lessons influencing students' non-cognitive physics related factors.

Overall, our results about these mindset shifts are an important basis for upcoming research on the effectiveness of mindset interventions and to answer the question which age group of students can benefit the most from interventions to foster a growth mindset in physics.

REFERENCES

Archer, L., Moote, J., & MacLeod, E. (2020). Learning that physics is 'not for me': Pedagogic work and the cultivation of habitus among advanced level physics students. *Journal of the Learning Sciences*, *29*(3), 347–384. https://doi.org/10.1080/10508406.2019.1707679

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit Theories of Intelligence Predict Achievement Across an Adolescent Transition: A Longitudinal Study and an Intervention. *Child Development*, *78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Cheng, Z., & Hau, K.-T. (2003). Are intelligence and personality changeable? Generality of Chinese students' beliefs across various personal attributes and age groups. *Personality and Individual Differences*, *34*(5), 731–748. https://doi.org/10.1016/S0191-8869(02)00030-2

Cimpian, A., Arce, H.-M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle Linguistic Cues Affect Children's Motivation. *Psychological Science*, *18*(4), 314–316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* (S. xiii, 195). Psychology Press.

Goldhorn, L., Wilhelm, T., Spatz, V., & Rehberg, J. (2020). Fixed und Growth Mindset: Selbstbilder von Schüler*innen in Physik. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*, *1*, 187–191. http://www.phydid.de/index.php/phydid-b/article/view/1030/1126

Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., & Levine, S. C. (2013). Parent Praise to 1- to 3-Year-Olds Predicts Children's Motivational Frameworks 5 Years Later. *Child Development*, *84*(5), 1526–1541. https://doi.org/10.1111/cdev.12064

Gunderson, E. A., Hamdan, N., Sorhagen, N. S., & D'Esterre, A. P. (2017). Who needs innate ability to succeed in math and literacy? Academic-domain-specific theories of intelligence about peers versus adults. *Developmental Psychology*, *53*(6), 1188–1205. https://doi.org/10.1037/dev0000282

Hoffmann, L., Häussler, P., & Lehrke, M. (1998). *Die IPN-Interessenstudie Physik*. IPN.

Jonsson, A.-C., Beach, D., Korp, H., & Erlandson, P. (2012). Teachers' implicit theories of intelligence: Influences from different disciplines and scientific theories. *European Journal of Teacher Education*, *35*(4), 387–400. https://doi.org/10.1080/02619768.2012.662636

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Limeri, L. B., Carter, N. T., Choe, J., Harper, H. G., Martin, H. R., Benton, A., & Dolan, E. L. (2020). Growing a growth mindset: Characterizing how and why undergraduate students' mindsets change. *International Journal of STEM Education*, *7*(1), 35. https://doi.org/10.1186/s40594-020-00227-2

Meyer, M., Cimpian, A., & Leslie, S.-J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00235

Ornek, F., Robinson, W. R., & Haugan, M. P. (2008). What Makes Physics Difficult? *International Journal of Environmental and Science Education*, *3*(1), 30–34.

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, C. M., … Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, *573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

THE INTEGRATION OF TEXT AND ILLUSTRATIONS IN BIOLOGY TEXTBOOKS USED IN GREEK LYCEUM

Georgios Ampatzidis[1] *and Anastasia Armeni*[2]

[1]University of Patras, Patras, Greece

[2]National and Kapodistrian University of Athens, Athens, Greece

Considering the fact that (a) science textbooks tend to rely heavily on illustrations and (b) textbooks determine largely what is taught and learned in science classrooms, this paper reports on exploring the illustrations included in four biology textbooks used in Greek lyceum (K10-12). Drawing on the Graphical Analysis Protocol (GAP), we coded the 581 illustrations identified in mutually exclusive categories concerning the text-illustrations integration – i.e. (a) the level of contiguity between illustrations and text, (b) the extent of in-text reference to illustrations, and (c) the captions' function regarding the connection of text and illustrations. The analysis of results suggests that (i) the majority of illustrations are contiguous with the relevant text in all four textbooks, (ii) the majority of illustrations are referenced in two of the textbooks we researched but there is only one referenced illustration in the other two, and (iii) the majority of illustrations in all textbooks have captions that are mostly used to identify or describe the illustrations.

Keywords: Graphical Representations, Curriculum, Science Education

INTRODUCTION

The communication of scientific and technical information has widely relied on illustrations since as early as the 1500s. Engineers of that time used notebooks and technical manuals to convey relevant information; to be more effective, as a general rule they were largely illustrated. In fact, sometimes text was not even included; when there was text, it only served to explain the illustrations. Following the invention of printing press in the 15th century, these illustrated books became available to large audiences and some historians of science support that their increased availability may have resulted in the technological advancements of 16th to 18th century (Hegarty, Carpenter, & Just, 1991).

In current times, images of science seem to play an important role in scientific inquiry itself, media reports of scientific news, and science textbooks (Slough, McTigue, Kim, & Jennings, 2010). Focusing on textbooks, we notice that the majority of science textbooks rely largely on illustrations (Ampatzidis & Armeni, 2021; Liu & Khine, 2016). Lee (2010) argues that among the printed materials used for the communication of science, textbooks are the ones that dedicate most of their space for illustrations. Mayer, Steinhoff, Bower, & Mars (1995) mention that nearly half of the page space in the textbooks they investigated is covered by illustrations. Moreover, Dimopoulos, Koulaidis, & Sklaveniti (2003) claim that the more recent a science textbook is, the more illustrations it contains, while Martins (2002) argues that that the diversity and representation of illustrations contained in science textbooks has been increased over the years. The increase of their use in recent years reflects the increasing importance of illustrations in education (Postigo & López-Manjón, 2019). Illustrations are important in science education because they seem to help students build understanding about scientific concepts which are inherently complex and abstract (Devetak & Vogrinc, 2013). It has been long appreciated that well-designed illustrations play an instrumental role in the teaching and learning of demanding scientific ideas (Khine, 2013).

Slough et al. (2010) argue that despite their importance in text comprehension, there are no consistent definitions of illustrations. Illustrations may be defined as polysemic and monosemic representations (Guo, Zhang, Wright, & McTigue, 2020) including a wide range of visual-spatial representations such as graphs, images, photographs, tables, diagrams, drawings and maps (Postigo & López-Manjón, 2019). They may be classified based on presentation (i.e. diagrams, maps, etc.) (Vekiri, 2002) on their function (i.e. diagrams, charts, graphs etc.) (Hegarty et al., 1991) or, following a mixed approach, in four types (i.e. pictures, pictorial diagrams, flow diagrams and mixed graphics) (Guo et al., 2020). Graphical demands of textbooks have been considerably explored concerning university education (Ampatzidis & Armeni, 2019; Bowen, Roth, & McGinn, 1999), secondary education (Liu & Treagust, 2013), and primary education (Liu & Khine, 2016; Slough et al., 2010).

Considering the above and drawing on previous research we have done with science textbooks used in a distance learning undergraduate course (Ampatzidis & Armeni, 2019) we decided to explore the illustrations included in biology textbooks used in Greek school. Our focus here is set in evaluating how well illustrations integrate with text in the Greek lyceum's biology textbooks; more specifically, the research questions (RQs) we address are:

- (RQ1) What is the level of contiguity between illustrations and text within Greek lyceum's biology textbooks?
- (RQ2) To what extent are illustrations referenced in the text within Greek lyceum's biology textbooks?
- (RQ3) In what manner are captions used to connect illustrations with text within Greek lyceum's biology textbooks?

METHODS

For this study, we investigated the 4 biology textbooks used in Greek lyceum (K10-K12) during the school year 2020-21:

- Biology for the 1st grade of Lyceum (Kastorinis, Kostaki-Apostolopoulou, Mparona-Mamali, Peraki, & Pialoglou, 2011): textbook-1.
- Biology for the 2nd grade of General Lyceum-General Education (Kapsalis, Bourmpouchakis, Peraki, & Salamastrakis, 2013): textbook-2.
- Biology for the 2nd and 3rd grade of General Lyceum (Adamantiadou et al., 2013): textbook-3.
- Biology for the 3rd grade of General Lyceum-Health Studies Specialization (Aleporou-Marinou, Argyrokastritis, Komitopoulou, Pialoglou, & Sgouritsa, 2013): textbook-4.

In order to answer our research questions, initially we counted the illustrations included in all four textbooks – the illustrations included in cover pages, appendices and assessment activities pages were excluded from our analysis. We identified 169 illustrations in textbook-1, 135 illustrations in textbook-2, 155 illustrations in textbook-3 and 122 illustrations in textbook-4. The 581 illustrations identified were coded in mutually exclusive categories formed drawing on the Graphical Analysis Protocol (GAP) proposed by Slough & McTigue, 2013). Table 1 shows our coding scheme after removing one empty category (i.e., 'unconnected' category) and merging categories with little content (i.e., 'proximal' and 'direct' categories were merged into the 'same page' category).

Both authors coded independently 150 (i.e., about 26%) randomly chosen graphics and the inter-rater agreement was about 98% for 'contiguity', 100% for 'indexical reference' and about 93% for 'captions' (see Table 1 for details on categories). The rest of the analysis was carried out by the first author.

Table 1. The coding scheme.

Contiguity	**Category**	**Description**
	Distal	Illustration and relevant text are in different pages (the reader needs to turn page)
	Facing	Illustration and relevant text are in facing pages
	Same page	Illustration and relevant text are in the same page
Indexical Reference	Referenced	Text does not reference the illustration
	Not referenced	Text references the illustration (e.g., 'See Figure 3.1')
Captions	No caption	There is no caption
	Identification	Caption identifies the target of the illustration but does not provide details
	Description	Caption provides a description of the illustration with details
	Engagement	Caption actively engages the readers (e.g., poses a question, asks them to read a specific part of the text, gives instructions)

RESULTS

Most illustrations are rather contiguous with the text in the textbooks investigated; illustrations positioned in the same page as the relevant text are more than distal and facing illustrations in all four textbooks (Figure 1). The biggest percentage of non-contiguous (i.e. distal and facing) illustrations are found in textbook-4 (41/122) while the smallest percentage are found in textbook-3 (20/155).

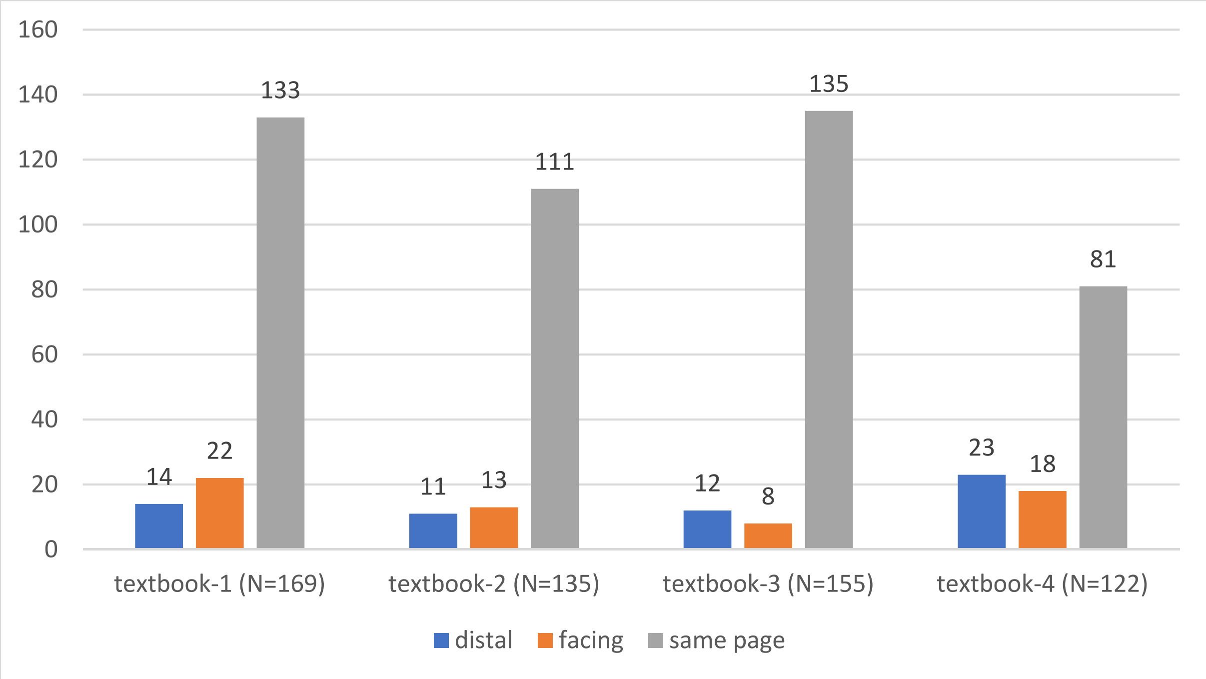


Figure 1. Frequencies of categories of illustrations regarding contiguity.

The majority of illustrations in two textbooks are referenced in the text. In parenthesis – e.g., '(Figure 1.13)' – or not – e.g. 'The excretory organs appear in Figure 6.1…' –the authors of textbook-1 and textbook-4 signal the reader when to view the relevant illustration. On the contrary, illustrations in textbook-3 are never referenced in the text while only 2/135 illustrations are referenced in the case of textbook-2 (Figure 2).

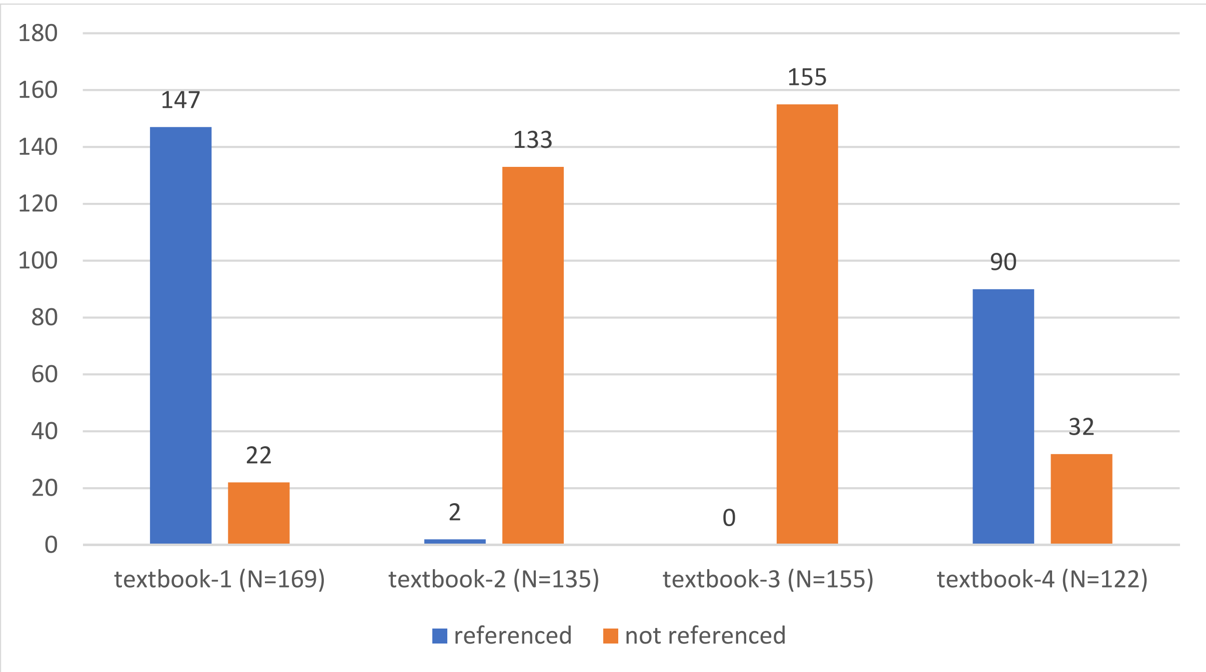


Figure 2. Frequencies of categories of illustrations regarding indexical reference.

Most illustrations in all textbooks have captions. When they exist, captions are used (a) to identify and (b) to describe the illustration in almost all cases. An example of (a) is the following: 'Capillary network that feeds a group of somatic cells' and an example of (b) is the following: '1. Cells are surrounded by the interstitial fluid which provides them with nutrients and in which they discharge waste products 2. Blood pressure at capillaries is high comparing to the interstitial fluid and as a result, small molecules of the plasma exit capillaries 3. Erythrocytes, as well as large molecules such as proteins, remain inside the capillaries.' There is only on caption that intends to engage the reader by giving instructions on performing a breast self-examination under a relevant illustration of textbook-1: 'Stand in front of the mirror and carefully examine your breasts for any change in their shape or colour. Lie down on your bed, having previously placed a pillow underneath from your right shoulder. Put your right hand under your head.'

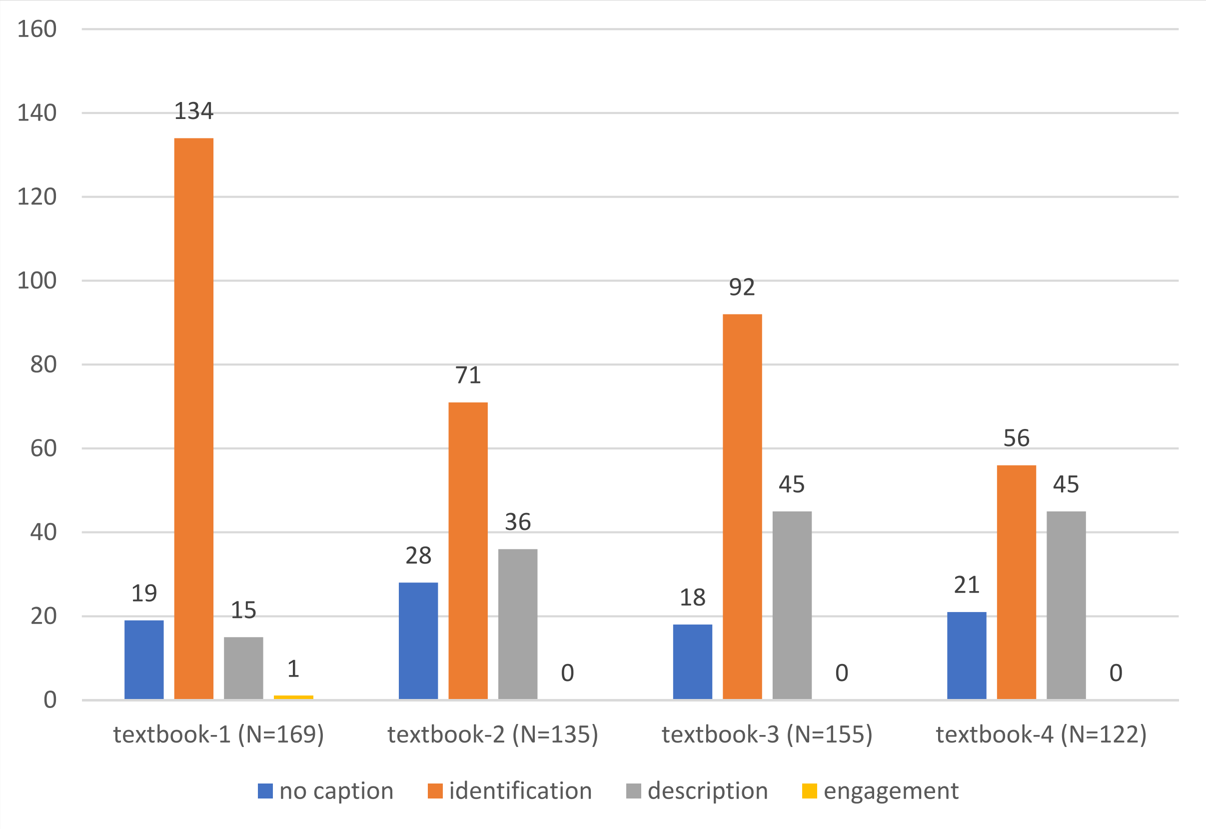


Figure 3. Frequencies of categories of illustrations regarding captions.

DISCUSSION

Mayer (2001) supports that the proximity between text and illustrations contributes to their effectiveness; students are believed to perform better at tasks when illustrations and relevant text are placed close to each other. Studies based on eye-tracking technology show that, in case of illustrated text, readers need to refer back and forth between text and illustrations many times. This observation seems to suggest that placing the two sources of information close to each other may help their shift of attention (Slough et al., 2010). It has been argued that design features which help students as they read the textbooks, such as proximity between text and illustrations, not only enhance the text-illustrations integration, but they also decrease cognitive load linked with instructional design (Nyachwaya & Gillaspie, 2016). It seems that the authors of the textbooks we investigated effectively achieve both goals in terms of proximity between text and illustrations, since in all four textbooks that concern our study most illustrations are positioned at the same page with relevant text.

Another feature that help students as they read the textbooks is indexing of illustrations within text (Nyachwaya & Gillaspie, 2016). References within running text is considered an effective way to signal the reader when to observe the illustration, arguably offering a form of guide of how the reader is supposed to integrate textual and visual information (Slough & McTigue, 2013). We note that the authors of the textbooks we researched form two distinct groups in regards to referring to illustrations within text: in two textbooks (textbook-1 and textbook-4) there are references to the majority of illustrations while in the rest (textbook-2 and textbook-3) there are no or almost none reference.

Finally, the use of the use of captions is also considered to help students as they read the textbooks. Captions are thought to be an efficient way to inform the reader when to refer to a relevant illustration and they provide a guide about how the readers may integrate the information conveyed by text and illustrations (Slough & McTigue, 2013). Our results suggest that most illustrations included in the textbooks we researched have captions that identify or describe what is depicted. Peeck (1993) notes that when captions simply signal readers to refer to illustrations, they do not do much to effectively support them in processing the visual information. He suggests that captions should ask the readers to do something with them – i.e. he notes that illustrations should have engaging captions. In our study we identified only one caption intending to engage the readers (textbook-1).

Exploring some aspects of the text-illustration integration in biology textbooks of Greek lyceum, we have made interesting observations regarding the effectiveness of illustrations. Although we have discussed how the effectiveness of illustrations is influenced by specific variables, we make no claim about whether and how the text-illustration integration results we discuss are linked with an improved or not students learning. Moreover, we do not assume our results to be generalizable, since our research concerns a limited number of textbooks originated in one country. However, this study follows previous research we have done asking similar research questions concerning textbooks used in a distance learning undergraduate science course (Ampatzidis & Armeni, 2019). We plan to further extend our study by investigating more science textbooks used in secondary and higher education.

REFERENCES

Adamantiadou, S., Georgatou, M., Giapitzakis, Ch., Lakka, L., Notaras, Δ., Florentin, N., … Chantikonti, O. (2013). *Biology for the 2nd and 3rd grade of General Lyceum*. Athens, Greece: CTI Diophantus.

Aleporou-Marinou, V., Argyrokastritis, A., Komitopoulou, A., Pialoglou, P., & Sgouritsa, V. (2013). *Biology for the 3rd grade of General Lyceum-Health Studies Specialization*. Athens, Greece: CTI Diophantus.

Ampatzidis, G., & Armeni, A. (2021). Analysis of illustrations concerning human reproduction included in Greek secondary education textbooks. *Mediterranean Journal of Education*, *1*(2), 46–54.

Ampatzidis, G., & Armeni, M. (2019). Graphic-text integration in distance education science textbooks. In A. Lionarakis, E. Manousou, S. Ioakimidou, M. Niari, A. Angeli, K. Sfakiotaki, & V. Koutzeklidou (Eds.), *Proceedings of the 10th International Conference in Open & Distance Learning: Shaping the Future of Education, Volume 10, Issue 2A* (pp. 117–124). Athens, Greece: Hellenic Open University and Hellenic Network of Open and Distance Education.

Bowen, G. M., Roth, W.-M., & McGinn, M. K. (1999). Interpretations of graphs by university biology students and practicing scientists: Toward a social practice view of scientific representation practices. *Journal of Research in Science Teaching*, *36*(9), 1020–1043.

Devetak, I., & Vogrinc, J. (2013). The Criteria for Evaluating the Quality of the Science Textbooks. In M. S. Khine (Ed.), *Critical Analysis of Science Textbooks: Evaluating instructional effectiveness* (pp. 3–15). Dordrecht, Netherlands: Springer.

Dimopoulos, K., Koulaidis, V., & Sklaveniti, S. (2003). Towards an Analysis of Visual Images in School Science Textbooks and Press Articles about Science and Technology. *Research in Science Education*, *33*(2), 189–216.

Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do You Get the Picture? A Meta-Analysis of the Effect of Graphics on Reading Comprehension. *AERA Open*, *6*(1), 1-20.

Hegarty, M., Carpenter, P. A., & Just, M. A. (1991). Diagrams in the comprehension of scientific texts. In R. Barr, M. L. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2). New York, NY: Longman.

Kapsalis, A., Bourmpouchakis, I. E., Peraki, V., & Salamastrakis, S. (2013). *Biology for the 2nd grade of General Lyceum-General Education*. Athens, Greece: CTI Diophantus.

Kastorinis, A., Kostaki-Apostolopoulou, M., Mparona-Mamali, F., Peraki, V., & Pialoglou, P. (2011). *Biology for the 1st grade of Lyceum*. Athens, Greece: CTI Diophantus.

Khine, M. S. (Ed.). (2013). *Critical Analysis of Science Textbooks*. Dordrecht, Netherlands: Springer.

Lee, V. R. (2010). Adaptations and Continuities in the Use and Design of Visual Representations in US Middle School Science Textbooks. *International Journal of Science Education*, *32*(8), 1099–1126.

Liu, Y., & Khine, M. S. (2016). Content Analysis of The Diagrammatic Representations of Primary Science Textbooks. *Eurasia Journal of Mathematics, Science and Technology Education*, *12*(8), 1937–1951.

Liu, Y., & Treagust, D. F. (2013). Content Analysis of Diagrams in Secondary School Science Textbooks. In M. S. Khine (Ed.), *Critical Analysis of Science Textbooks: Evaluating Instructional Effectiveness* (pp. 287–300). Dordrecht, Netherlands: Springer.

Martins, I. (2002). Visual imagery in school science texts. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 73–90). Mahwah, NJ: Lawrence Erlbaum.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, NY: Cambridge University Press.

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, *43*(1), 31–41.

Nyachwaya, J. M., & Gillaspie, M. (2016). Features of representations in general chemistry textbooks: A peek through the lens of the cognitive load theory. *Chemistry Education Research and Practice*, *17*(1), 58–71.

Peeck, J. (1993). Increasing picture effects in learning from illustrated text. *Learning and Instruction*, *3*(3), 227–238.

Postigo, Y., & López-Manjón, A. (2019). Images in biology: Are instructional criteria used in textbook image design? *International Journal of Science Education*, *41*(2), 210–229.

Slough, S. W., & McTigue, E. M. (2013). Development of the Graphical Analysis Protocol (GAP) for Eliciting the Graphical Demands of Science Textbooks. In M. S. Khine (Ed.), *Critical Analysis of Science Textbooks* (pp. 17–30). Dordrecht, Netherlands: Springer.

Slough, S. W., McTigue, E. M., Kim, S., & Jennings, S. K. (2010). Science Textbooks' Use of Graphical Representation: A Descriptive Analysis of Four Sixth Grade Science Texts. *Reading Psychology*, *31*(3), 301–325.

Vekiri, I. (2002). What Is the Value of Graphical Displays in Learning? *Educational Psychology Review*, *14*(3), 261–312.

DEVELOPMENT OF A SCIENCE TEACHING AND LEARNING AVERSION SCALE (STLAS) AND CLASSIFICATION OF UNDERGRADUATE STUDENTS

Shotaro Naganuma[1]

[1]Kyushu University, Fukuoka, Japan

This study has two purposes: (i) to examine the reliability and validity of a newly developed scale, 'Science Teaching and Learning Aversion Scale (STLAS)' (Study 1), and (ii) to classify undergraduate students according to their attitudes and extract their characteristics (Study 2). In Study 1, we developed 28 items based on data collected from 22 freshmen and sophomores at a public university. A total of 320 undergraduate students completed the survey. The results indicated that STLAS contained seven factors (performing calculations, applying abstract concepts, failure in experiments, grades and understanding, subjects for study, life irrelevance, and teachers' characteristics); of them, four (performing calculations, applying abstract concepts, failure in experiments, and life irrelevance) demonstrated reliability and validity. In Study 2, we identified five profiles with STLAS for 1075 undergraduate students based on model fit and interpretability: extreme (14.0%), mild (28.7%), neutral (26.4%), science OK (23.4%), and science lover (7.5%). We observed a similar pattern in the scores on two types of needs for cognition (in general and science). This observation generated two hypotheses about the relationship between learners' attitudes towards science and general academics that require further examination. Subsequent analysis revealed three topics: 1) science lovers and science OKs are more interested in these six topics than mild and extreme anti; 2) for many topics, the degree of interest of neutral and mild anti was almost similar; and 3) the possibility of attracting the interest of extreme anti by introducing these six topics in science classes seemed quite low. These findings indicate a need to reconsider the means to pique the interest of the neutral, mild, and extreme anti groups.

Keywords: attitudes towards science; latent profile analysis; ANOVA

INTRODUCTION

The lack of interest in science has been viewed as a serious issue for several decades (Osbourne, et al. 2003; Vedder-Weiss & Fortus, 2011; Tytler, 2014, Tytler & Osbourne, 2012). Several intervention programs have been developed, including problem-based learning, hands-on activities, and summer camps (Harackiewicz et al. 2016; van den Hurk et al., 2019). Nevertheless, as demonstrated by international survey results (e.g., Mullis et al., 2019), the solution to this issue remains elusive. The author believes that a more detailed examination of negative attitudes towards science may lead to a solution. Recently, several studies have been conducted to categorise attitudes towards science using a person-centred approach (Perez et al. 2018; Radišić et al., 2020; Rangel et al., 2020), but these studies used highly granular items. The lack of an appropriate scale to measure negative attitudes towards science on a fine-grained scale and studies using these scales to classify learner types has resulted in limited insights into whether an intervention is effective for all types of students.

Therefore, two studies were conducted. Study 1 developed the 'Science Teaching and Learning Aversion Scale (STLAS)'. We aimed to examine the reliability and validity of this scale. Study 2 aimed to categorise students based on their attitudes towards school science and attempted to reveal the features of each profile. Specifically, we measured attitudes towards science teaching and learning using STLAS to classify students. We then conducted a latent profile analysis, a person-oriented approach, to identify the participant's profile. To examine the characteristics of each group, an analysis of variance (ANOVA) was used to examine associations among the resulting profiles and three outcomes: needs for cognition in general, needs for cognition in science, and students' interest in studying science-related topics.

METHOD (STUDY 1)

Study 1 aimed to examine the reliability and validity of a newly developed scale: the STLAS.

1. Measurements

a. Science Teaching and Learning Aversion Scale (STLAS)

First, as a preliminary survey for item development, an interview survey was conducted with 22 first- and second-year undergraduate students. All the participants were university students at a public university who had previously experienced disliking science. Items, developed using the information collected, were classified into seven groups (performing calculations, applying abstract concepts, failure in experiments, grades and understanding, subjects for study, life irrelevance, and teachers' characteristics).

Next, items were deleted or added, such that each group would have four items each, with a total of 28 items (Table 1). The instructions were as follows: 'We would like to ask you about the science classes you took in high school (including "Basic physics" and "Chemistry"). To what extent do you agree with each of the following statements?' For each item, respondents were asked to choose from five options: strongly disagree (1), disagree (2), cannot say either (3), agree (4), and strongly agree (5). The numbers in parentheses are the scores assigned for quantitative analysis. In this study, we defined 'science teaching and learning aversion' as 'feelings of aversion, hate, or dislike to various events in school science teaching and learning situations.

b. Enjoyment of science

We used the 'Enjoyment of Science' indicator from the Programme for International Student Assessment survey (National Institute for Educational Policy Research, 2016, for the Japanese version) for validity examination. We assumed that the higher the 'Enjoyment of Science' score, the higher the score of the ATLAS. An example item is 'I like reading about science'. Although this indicator originally contains five items, one item ('I am happy working on <broad science> topics') was excluded because humanities students tend to have limited opportunities to solve scientific problems. Students were asked to choose from the following four options: strongly disagree (1), disagree (2), agree (3), and strongly agree (4). The numbers in parentheses are the assigned scores. The mean value for the four items was used as a score for subsequent analysis.

2. Participants

An online survey was administered in August 2020 among university undergraduates (grades 1–4) registered in Cross Marketing Inc., a Japanese Internet research company. The data of 320 students (96 men, 220 women, and 4 others) were analysed after excluding those who continuously selected the same number, those who showed missing values, and those who gave inappropriate answers to the trap question (i.e., 'Select the leftmost option for this item'). The average age of the respondents at the time of the survey was 20.35 years (SD = 1.51).

3. Data analysis

The reliability of the STLAS was examined using Cronbach's alpha coefficients. First, a factor analysis was used to examine the validity. Then, the correlation coefficient between each factor and the indicator of 'Enjoyment of Science' was calculated. If the STLAS is valid, there should be a negative correlation. We used the free statistical software, HAD (Shimizu, 2016).

RESULTS (STUDY 1)

1. Descriptive statistics

First, the mean and standard deviation values for each item were calculated (Table 1). There were no ceiling or floor effects.

2. Factor analysis

An exploratory factor analysis (maximum likelihood method and Promax rotation) was conducted on the data for all 28 items. Based on the decline of the eigenvalues, the Kaiser criterion, the minimum average partial correlation, and interpretability, a seven-factor structure was determined to be appropriate. However, during this process, one reverse item that behaved differently from our assumption ('hated conducting experiments to understand abstract concepts') and one item with a factor load below 0.4 ('hated being asked to memorise contents without understanding them') were extracted. Therefore, these items were excluded, and the 26 items were again subjected to exploratory factor analysis, but the model was not optimised. The data were then reanalysed using the least squares method. Finally, the initially assumed seven-factor structure was extracted (Table 1). The cumulative factor contribution rate was 69.88%. The mean value of each factor was used for the analysis.

3. Reliability

The alpha coefficients were estimated as a measure of the reliability of each scale score; for six sub-concepts, the alpha was .75 or higher (Table 2). Only 'study as a subject' had a low alpha coefficient of .58.

4. Validity

To examine the validity, we calculated the correlation coefficients between the seven factors in STLAS and 'Enjoyment of Science' (Table 2). There were significant moderate negative correlations between 'performing calculations', 'applying abstract concepts', and 'life irrelevance' and 'Enjoyment of Science', and significant small negative correlations between 'failure in experiments' and 'subjects for study' and 'Enjoyment of Science'. There were no significant correlations between 'grades and understanding' and 'teachers' characteristics' and 'Enjoyment of Science'.

Table 1. Descriptive Statistics.

	Items	M	SD	F1	F2	F3	F4	F5	F6	F7
F1	hated solving complex calculation problems	3.52	1.30	**0.89**	0.01	-0.05	0.06	0.01	-0.01	-0.04
	hated solving a lot of calculation problems	3.27	1.26	**0.82**	-0.08	0.04	-0.01	0.03	0.02	0.02
	liked classes that were conducted with mathematical formulae*	3.15	1.27	**0.61**	-0.05	-0.06	-0.07	-0.03	0.02	0.02
	hated taking lectures about calculation problems	2.85	1.29	**0.58**	0.18	0.09	-0.04	0.01	-0.06	0.00
F2	hated learning invisible concepts	2.99	1.20	-0.06	**0.91**	0.04	-0.08	0.03	0.01	-0.03
	hated learning concepts that were hard to imagine	3.27	1.12	0.00	**0.78**	0.05	0.08	-0.05	-0.04	0.04
	hated learning concepts that could not be replaced by the things around me	3.05	1.22	0.18	**0.67**	0.02	-0.03	0.04	0.04	-0.07
	liked learning abstract concepts*	2.90	1.05	-0.04	**0.44**	-0.19	0.01	-0.10	0.04	0.05
F3	hated failing to manipulate experiments	3.01	1.19	0.00	-0.09	**0.90**	0.01	-0.11	0.11	-0.06
	hated failing to obtain an expected result from an experiment	2.98	1.20	0.07	-0.05	**0.80**	0.02	0.03	-0.08	0.01
	hated not knowing the purpose of experiments	3.22	1.21	-0.15	0.08	**0.47**	0.03	0.09	-0.02	0.14
F4	hated getting low scores on tests	3.71	1.17	-0.13	0.04	0.02	**0.94**	-0.14	-0.05	-0.10
	hated getting poor grades compared to classmates	3.19	1.32	-0.02	-0.04	0.02	**0.76**	0.07	-0.02	0.03
	hated being unable to understand what your classmates understood	3.66	1.17	-0.01	0.02	-0.03	**0.74**	0.08	0.02	-0.01
	hated getting low test scores despite studying hard	3.58	1.17	0.15	-0.07	0.02	**0.68**	-0.07	0.04	0.05
F5	hated studying while worrying about test scores	3.22	1.21	-0.01	0.06	-0.04	0.11	**0.73**	-0.01	0.04
	liked making efforts only to improve your grades	2.81	1.16	0.02	-0.09	-0.01	-0.13	**0.54**	-0.03	-0.09
	hated studying for entrance exams as a subject	3.19	1.17	-0.02	-0.07	0.02	0.11	**0.44**	0.13	0.11
F6	hated learning content that appeared irrelevant to my future	3.23	1.20	-0.03	-0.02	0.01	-0.02	0.05	**0.97**	-0.07
	hated learning content that appeared useless to learn	3.22	1.22	0.02	-0.07	0.02	-0.05	-0.05	**0.94**	0.05
	hated learning content seemingly unrelated to my daily life	3.22	1.17	0.01	0.04	0.03	-0.01	0.14	**0.81**	-0.10
	hated learning content unworthy of learning	3.34	1.14	0.02	0.17	-0.05	0.07	-0.17	**0.72**	0.13
F7	hated having teachers who did not answer your questions properly	3.08	1.21	-0.03	-0.09	0.06	-0.07	0.02	-0.06	**0.87**
	hated being taught what teachers did not understand	3.16	1.22	-0.13	0.05	0.01	-0.05	-0.01	-0.01	**0.78**
	hated receiving lectures wherein teachers did not teach solid fundamentals for solving problems	3.28	1.18	0.06	0.08	-0.01	0.03	-0.09	0.04	**0.75**
	hated having teachers whose teaching style was difficult to understand	3.59	1.16	0.14	-0.02	-0.05	0.05	0.01	0.02	**0.71**

Note. F1 = Performing calculations, F2 = Applying abstract concepts, F3 = Failure in experiments, F4 = Poor grades and understanding, F5 = Study as a subject, F6 = Life irrelevance, and F7 = Teacher's characteristics

Table 2. Correlation Coefficient

	M	SD	α	F1	F2	F3	F4	F5	F6	F7	ES
F1	3.20	1.03	0.81	1.0							-.444**
F2	3.05	0.90	0.78	.459**	1						-.421**
F3	3.07	0.98	0.75	.231**	.260**	1					-.197**
F4	3.54	1.01	0.85	.185**	.154**	.417**	1				-0.019
F5	3.07	0.87	0.58	.234**	.218**	.221**	.388**	1			-.114*
F6	3.26	1.07	0.92	.418**	.528**	.325**	.305**	.345**	1		-.434**
F7	3.28	0.99	0.85	.128*	.175**	.257**	.358**	.194**	.337**	1.0	-0.019

Note. ES = Enjoyment of Science; ** p < .01; F1 = Performing calculations, F2 = Applying abstract concepts, F3 = Failure in experiments, F4 = Poor grades and understanding, F5 = Study as a subject, F6 = Life irrelevance, and F7 = Teacher's characteristics

METHOD (STUDY 2)

Study 2 aimed to classify learners according to their attitudes and extract their characteristics.

2.1. Measures

a. Science Teaching and Learning Aversion Scale (STLAS)

In Study 1, out of seven factors, we could not obtain reliability evidence for one factor ('study as a subject') and validity evidence for two factors ('grades and understanding' and 'teachers' characteristics'). Therefore, we used only four factors (performing calculations, applying abstract concepts, failure in experiments, and life irrelevance) in Study 2.

b. Need for cognition

Need for cognition is defined as 'an individual's tendency to engage in and enjoy effortful cognitive endeavours' (Cacippo et al., 1996, p.197). The sample item is 'I tend to set goals that can be accomplished only by expending considerable mental effort'. This scale was originally developed by Cacippo and Petty (1982), and the 15-item short Japanese version by Kouyama and Fujihara (1991) was used in this study. For each item, respondents were asked to answer on a five-point scale: strongly disagree (1), disagree (2), cannot say either (3), agree (4), and strongly agree (5).

c. Need for cognition in science

Need for cognition in science is defined as 'an individual's tendency to engage in a series of problem-solving through observation and experiments and enjoy it' (Unzai and Nakamura, 2018, p.302). This study aimed to not only classify people but also obtain insights for interventions. Therefore, two measures were useful: science-specific (need for cognition in science) and domain-general (need for cognition). The sample item is 'It is enjoyable to explain natural phenomena with scientific knowledge'. This scale was developed by Unzai and

Nakamura (2018) in Japanese as an embodiment of the Need for Cognition Scale by Kouyama and Fujihara (1991) in the context of science. Ten items were prepared for primary and junior high school students. Thus, to adopt it for undergraduate students in this study, minor corrections and exclusion of a few inappropriate items were conducted. For each item, respondents were asked to answer on a five-point scale: strongly disagree (1), somewhat disagree (2), cannot say either (3), somewhat agree (4), and strongly agree (5).

d. Interest in studying science-related topics

In Relevance of Science Education, an international survey, 108 items were used to collect students' interest in science (Schreiner & Sjoberg, 2004). In this study, we included 48 items that seemed related to science but not conventional ones (e.g., 'medical use of plants', 'how computers work'). For each item, respondents were asked to answer on a four-point scale: not interested (1), less interested (2), somewhat interested (3), and interested (4). After factor analysis, we identified six factors, including environmental protection and health.

2.2. Participants

A total of 1075 undergraduate students (825 women, 243 men, and 7 unknown) responded to the survey in September 2020 who had registered at the Cross Marketing, an Internet research system in Japan. We excluded those who continuously selected the same number, who showed missing values, and who gave inappropriate answers to the trap question (e.g., 'Select the leftmost option for this item'). The average age of the respondents at the time of the survey was 20.37 years (SD = 1.38).

2.3. Data analysis

A latent profile analysis was used to identify several attitudinal profiles. ANOVA was then used to examine how participants' profile membership was related to their need for cognition, need-for-cognition in science, and interest in studying science-related topics. Figure 1 shows the analysis model.

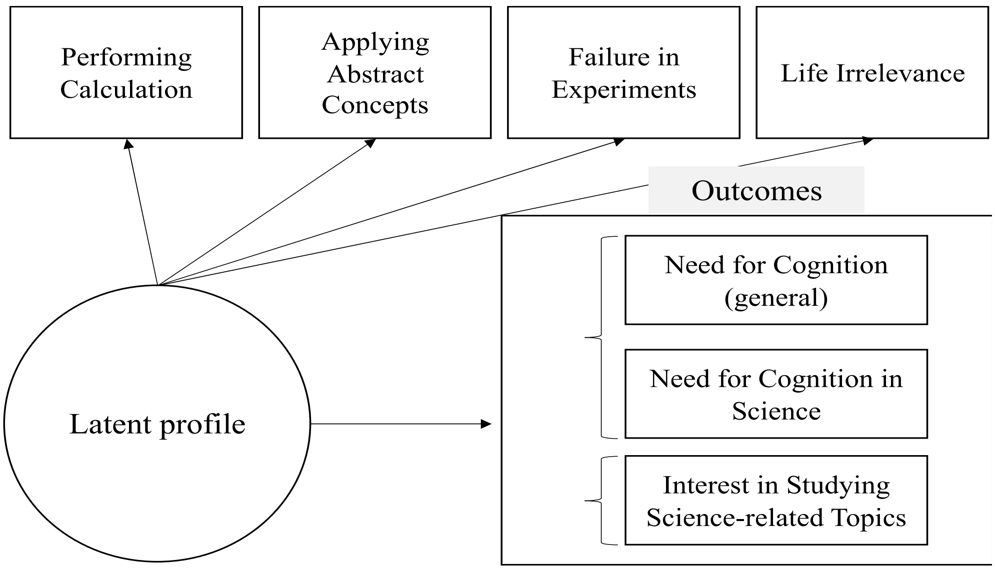


Figure 1. Analysis Model.

RESULTS (STUDY 2)

3.1. Structure of profiles

The structure of profile fit indices for potential profile solutions revealed that a five-profile solution was the best fit for the data from the current sample (Table 3).

Figure 2 demonstrates the five profiles. The first profile was characterised by the highest scores for all four factors; hence, we labelled this profile as 'extreme anti' to reflect their negative attitude towards all aspects. The second profile was characterised by a moderately negative attitude towards the four aspects. This profile was the largest. We labelled this as 'mild anti'. The third profile was characterised by neutral attitudes towards the four aspects. We labelled this profile as 'neutral'. The fourth profile was characterised by its moderately positive attitudes towards the four factors. Therefore, we labelled this as 'science OK'. Scores on all factors in the fifth profile are the smallest, implying that they generally have an extremely positive attitude towards school science. We labelled this profile as 'science lovers.

Table 3. Fit Indices for Different Latent Profile Solutions

# of profiles	BIC	ABIC	p for LMR	p for Bootstrap
1	12131.011	12105.601	N/A	N/A
2	11629.918	11588.627	0	0
3	11498.379	11441.208	0.043	0
4	11446.76	11373.707	0	0
5	11264.597	11175.664	0	0
6	11440.189	11335.374	0.3869	0
7	11463.757	11343.061	0.1914	0.5
8	11263.038	11126.462	0.1411	0
9	11155.228	11002.771	0.0004	0
10	11295.342	11127.004	0.2344	0.6667

Note. The five-profile solution was selected as the best-fitting solution.
BIC: Bayesian information criterion;
ABIC: adjusted Bayesian information criterion;
LMR: Lo–Mendell–Rubin adjusted likelihood ratio test.

3.2. Profile membership and two needs-for-cognition

The results from ANOVA in Figure 3 indicate that for both need for cognition (general) and need for cognition in science, the highest mean values are, in order, science lovers, science OK, neutral, mild anti, and extreme anti. Statistically significant differences were found among all the groups.

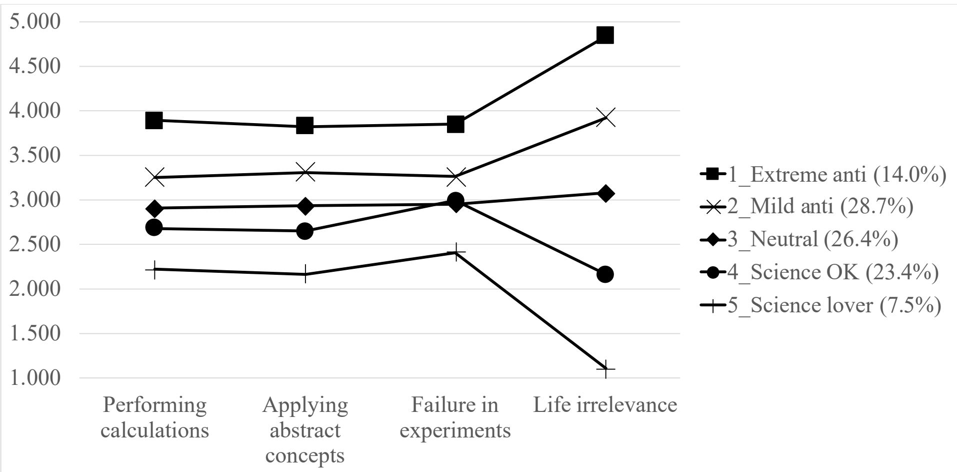


Figure 2. School Science Attitudinal Profile.

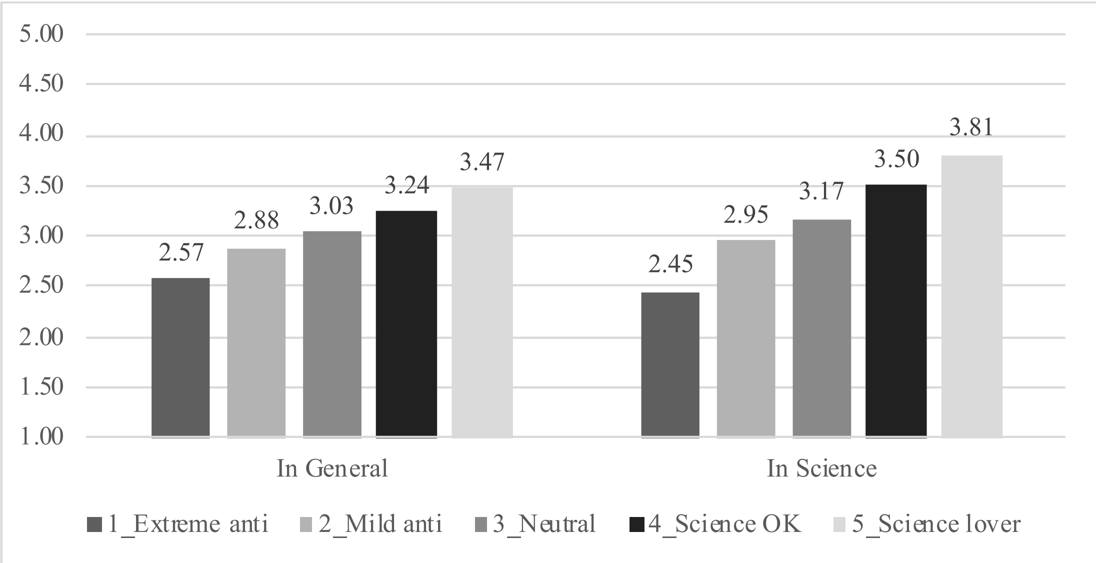


Figure 3. ANOVA: Profile and Need-for-Cognition.

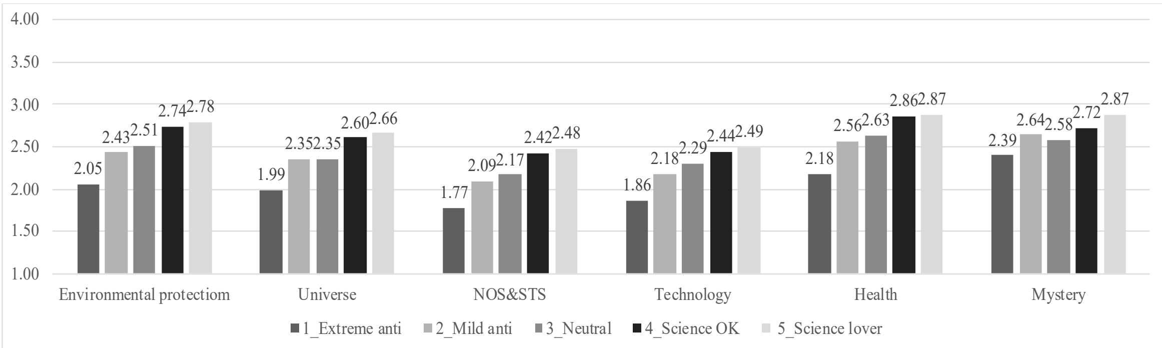


Figure 4. ANOVA: Profile and Interest in Studying Science-related Topics

3.3. Profile membership and interest in science-related topics

We performed an ANOVA to test the differences between participants' interest in studying science-related topics for the participants' profiles. The results in Figure 4 show that the mean values were larger in the order of science lovers, science OK, neutral, mild anti, and extreme anti. Mild anti did not differ significantly from the neutral on the remaining six topics except for mystery, and mild anti did not differ significantly from the science lover and science OK groups on mystery. There was no significant difference between the mild anti and neutral groups for the six topics, or between the science lovers and the science OK group for mystery. However, extreme anti was significantly different from the mean values of all profiles except for mystery. There was no significant difference between the extreme anti and neutral groups.

4. DISCUSSION

Study 1 examined the reliability and validity of the newly-developed ATLAS, which focused on students' negative attitudes towards school science on a fine scale. We developed a 28-item scale and confirmed the reliability and validity of the following four factors out of seven: performing calculations, applying abstract concepts, failure in experiments, and life irrelevance. We could not obtain reliability evidence for 'study as a subject'. Any exclusion of items under 'study as a subject' did not improve the alpha coefficient, thus requiring further examination. Likewise, validity evidence for 'grades and understanding' and 'teachers' characteristics' was not obtained. The reason might be that the negative attitudes towards these two factors are typical among participants and do not directly lead to their response to the 'Enjoyment of Science' indicator. Therefore, additional evidence should be collected to establish the validity of these two factors.

Study 2 categorised participants into five profiles with STLAS: science lovers, science OK, neutral, mild anti, and extreme anti. Interestingly, the lines of the five profiles did not intersect, except for one point of failure in the experiments. In other words, unlike our initial assumptions, we did not detect any profile that scored high on one factor but low on another factor. Additionally, when compared to the other three factors, life irrelevance was characterised by a substantially larger difference in scores between profiles. This result suggests that emphasising the relevance of learning content to real life would be highly effective in attracting the interest of the mild anti and extreme anti groups. A similar pattern of scores on the two types of need for cognition (in general and in science) suggests two interesting hypotheses that need further examination: 1) Attitudes towards science may influence their attitude towards academics in general, and 2) attitudes towards science may be influenced by attitudes towards academics in general. If the first hypothesis is supported, issues related to students' attitude towards science must be promptly addressed since it affects their attitudes towards other subjects beyond the context of science. If the second hypothesis is supported, issues related to students' attitude towards science are hardly resolved only in the context of science and require taking other subjects into consideration. The results of the analysis between the profiles and interest in studying science-related topics revealed the following three topics: 1) Science lovers and science Oks are more interested in these six topics than mild anti and extreme anti; 2) for many topics, the degree of interest of neutral and mild anti was almost similar; and 3) the possibility of attracting the interest of extreme anti by introducing these six topics in science classes seemed quite low. While science, technology, engineering, and mathematics (STEM) education is expected to increase students' interest (Martín-Páez et al., 2019), this study raises concerns about the effectiveness of STEM education interventions for extreme anti and mild anti profiles. Given the above results, there are two possible future research directions. First, the four factors in STLAS used in Studies 1 and 2 must get attention from program developers. For example, if students find it difficult to perform calculations, an instruction focusing on improving their calculation skills may be effective in improving their attitude towards science. Another approach includes involving other subjects to solve attitudes towards science issues. The results suggest that science-related topics (e.g., environmental protection, Technology) are insufficient; hence, we may need to transcend beyond the context of science and include more distant subjects, such as language arts, history, politics, and so on. These two directions need to be examined in more detail in the future.

ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI [grant number 20K14115]. We would like to thank Editage (www.editage.com) for English language editing.

REFERENCES

Cacioppo, J. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116–131.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197–253

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences, 3*(2), 220–227.

Kouyama, T. & Fujihara, T. (1991). A basic study of the need for cognition scale. *Japan Journal of Social Psychology, 6*(3), 184–192.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D., & Fishbein, B. (2020) TIMSS 2019 International Results in Mathematics and Science. Boston, USA: TIMSS & PIRLS International Study Center.

Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education, 103*, 799–822.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25,* 1049–1079.

Perez, T., Wormington, S. V., Barger, M. M., Schwartz-Bloom, R. D., Lee, Y., & Linnenbrink-Garcia, L. (2019). Science expectancy, value, and cost profiles and their proximal and distal relations to undergraduate science, technology, engineering, and math persistence. *Science Education, 103*(2), 264–286.

Radišić, J., Selleri, P., Carugati, F., & Baucal, A. (2021). Are students in Italy really disinterested in science? A person-centered approach using the PISA 2015 data. *Science Education, 105*(2), 438–468.

Rangel, V. S., Vaval, L., & Bowers, A. (2020). Investigating underrepresented and first-generation college students' science and math motivational beliefs: A nationally representative study using latent profile analysis. *Science Education, 104*(6), 1041–1070.

Schreiner, C. & Sjoberg, S. (2004). Sowing the seeds of ROSE Background, rationale, questionnaire development and data collection for ROSE (The Relevance of Science Education) – a comparative study of students' views of science and science education. *Acta Didactica, 4*, Oslo, Dept. of Teacher Education and School Development, University of Oslo.

Shimizu, H. (2016). An introduction to the statistical free software HAD: Suggestions to improve teaching, learning and practice data analysis. *Journal of Media, Information and Communication*, *1*, 59–73. (in Japanese)

Tytler, R. (2014). Attitudes, identity, and aspirations toward science. *Handbook of Research on Science Education* (pp. 82–103). New York: Routlege.

Tytler, R., & Osborne, J. (2012). Student attitudes and aspirations towards science. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), *Second International Handbook of Science Education* (pp. 597–625). Dordrecht: Springer Netherlands.

Unzai, H. & Nakamura, D. (2018). Developing a need-for-cognition-in-science Scale. *Japan Journal of Science Education, 42*(4), 301–313. (in Japanese)

van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education, 41*(2), 150–164.

Vedder-Weiss, D. & Fortus, D. (2011). Adolescents' declining motivation to learn science: Inevitable or not? *Journal of Research in Science Teaching, 48*, 199–216.