Show Me The Learning
Adelaide, Nov 27-30, 2016

All roads lead to Rome: Tracking students' affect as they overcome misconceptions

Gregor Kennedy
Melbourne Centre for the Study of Higher Education
University of Melbourne

Jason M. Lodge
Melbourne Centre for the Study of Higher Education
University of Melbourne

Helping students to overcome misconceptions is a complex problem in digital learning environments in which students need to monitor their own progress and self-regulate their own learning. This is particularly so in flexible, discovery-based environments that have been criticised for the lack of support and structure provided to students. Emerging evidence suggests that discovery-based environments might be ineffective due to students becoming confused, frustrated or bored. In the study reported here, we examined the affective experience of students as they worked to overcome a common misconception in a discovery-based environment. While the results suggest that students experience a range of emotions, they all successfully overcame their initial misconception. Implications for the investigation of student affect in discovery-based environments and the design of these environments are also discussed.

Keywords: misconceptions, confusion, discovery-based learning, digital learning environments

Introduction

Misconceptions are common in many disciplinary areas. Contrary to traditional assumptions that students enter educational situations as blank slates, students in fact often have extensive intuitive notions of concepts or ideas. These 'folk' notions are also often not an accurate representation of the concept as it is understood scientifically. Misconceptions are particularly evident in areas such as physics (e.g. Brown, 1992) and psychology (e.g. Kuhle, Barber, & Bristol, 2009), where students' experiences in the world and interactions with other people do not necessarily match a more sophisticated understanding of the physical or psychological world as uncovered by science. In this paper, we describe a study that examined how students overcame misconceptions in a digital, discovery-based learning environment. The purpose of this study was to uncover the different experiences students have as they overcome scientific misconceptions when given flexibility to explore a relatively unstructured digital learning task.

Discovery-based learning environments have been the focus of much discussion over several decades (see Bruner, 1961). On the one hand, there is evidence to suggest that learning environments with minimal guidance that encourage student exploration can be effective in helping students develop strategies for enhanced conceptual understanding (De Jong & Van Joolingen, 1998). On the other hand, an argument made by some in the educational research community is that lightly scaffolded learning does not work and that students need explicit guidance to effectively work through learning tasks (Alfieri, Brooks, Aldrich & Tenenbaum, 2011). According to this view, giving students flexibility to determine their own path through a learning task ostensibly creates extraneous cognitive load that hampers learning (Kirschner, Sweller & Clark, 2006). So while there is evidence to suggest that discovery-based learning environments can pose problems for learning generally (see also Hattie, 2009), there is some evidence to suggest that they can be effective for enhancing conceptual understanding (Dalgarno, Kennedy & Bennett, 2014).

Despite the uncertainty around the prospect of using discovery-based environments, there is evidence to suggest these environments can be effective under particular circumstances. Dalgarno et al. (2014) found that a discovery-based environment assisted students to learn about blood alcohol concentration or global warming in comparison to tutorial versions of the same lessons. The caveat in this instance was that students needed to be systematic in their approach to navigating the lesson. Following on from this we (Lodge & Kennedy, 2015) have found that confidence, confusion and perceived difficulty also contributed to the strategies that students employ when working through discovery-based lessons in digital environments.

Previous research has also indicated that, aside from their approach or strategy, students' affective experiences during their learning can serve as important predictors of their progress in overcoming misconceptions. Previous research (e.g. D'Mello & Graesser, 2014; D'Mello, Lehman, Pekrun & Graesser, 2014) has suggested that students in both discovery-based and other learning environments experience cognitive disequilibrium as a result of being exposed to information contradicting their intuitive conceptions. This disequilibrium often results in the epistemic emotion (i.e. an affective response) of confusion. Despite confusion intuitively being seen as an indicator of ineffectiveness in discovery-based learning environments, there is evidence to suggest that confusion can be an essential part of an effective student learning process. While there are certainly times when confusion can impede learning, there are circumstances in which it can be particularly useful for students to achieve meaningful conceptual change (D'Mello, Lehman, et al., 2014).

Lehman, D'Mello and Graesser (2012) found that in "breakdown" scenarios (i.e. problem-based exercises where the task is to determine why a system or machine is not functioning) presented in digital environments, conflicting information and false feedback can all reliably generate confusion. If not adequately resolved, this confusion can lead to the further emotional responses of frustration and boredom. It is likely, in the circumstances where a student is persistently confused to the point of boredom or frustration, that the original misconception will remain or even be reinforced, rather than corrected. This has been referred to as a 'backfire effect' (Trevors, Muis, Pekrun, Sinatra & Winne, 2016). This confusion, boredom and frustration may explain why "pure" discovery-based environments are seen as ineffective (e.g. Mayer, 2004); if students are not able to successfully navigate the environment, they are likely to experience persistent confusion leading to frustration, boredom and eventually this may cause them to give up. However, when confusion is effectively resolved, it has the potential to lead to deeper learning, particularly of complex concepts (D'Mello & Graesser, 2014). What this suggests is that the confusion that is commonly seen as a negative aspect of discovery-based environments can in fact be both productive and necessary for conceptual change, as long as it can be resolved in a timely manner.

The overall purpose of the current study was to examine the affective experiences of participants using a digital learning task that was designed to help students overcome a scientific misconception. Previous research in this area, including our own (e.g. Lodge & Kennedy, 2015), has suggested that prior knowledge is a critical factor in determining how students approach a discovery-based digital learning environment. In order to control for this variability, we have focussed in this study on a scientific concept that is largely misconceived by novices in the area. As detailed in the method section below, the concept students were asked to explore was the relationship between the size of a star and its lifespan. Intuitively larger stars should have longer lifespans but, due mostly to the effects of gravity, the opposite is true (see Schwarzschild, 1958). Using a misconception as a starting point has allowed us to explore a scenario in which we were able to control for prior knowledge to allow greater emphasis on examining students' affective trajectory though the task. In effect we wanted all students to have a misconception and limited knowledge so that they predominantly start from a similar point (i.e. believe that larger stars live longer). The purpose of the study was thus to delve into the ways in which students move from this point towards resolution of the misconception (i.e. come to understand that smaller stars have longer lifespans).

Intrinsic to the design of relatively open, discovery-based learning environments, is the ability of students to take a range of learning pathways in order to come to their own understanding of the material. As indicated above, a concern with these environments is that these pathways, and students' learning processes that are aligned with these pathways, do not always lead to productive learning. While there are a range of ways that students can approach open-ended, discovery based tasks – that is, in fact the point – in this study we were interested in whether there were patterns associated with particular affective experiences, and whether these could be related to the successful resolution of a fundamental misconception. Thus a central aim of this research was to determine how the affective experience of students changed as they negotiated a discovery-based learning task, and to see how this was related to their (hopeful) resolution of a misconception. The approach used in this study adds to the research to date by examining the changes in affective experience throughout the course of conceptual change. We then compare and contrast this to the global, overall impression that participants have of the session, which is the most common means of assessing the experiences students have of completing a learning task. Having a better understanding of the affective trajectories of students will allow educators and designers to not only modify the design of learning environments to accommodate students' learning needs, but it could also form the basis of interventions to support students when their intuitive notions are challenged, thereby leading them to meaningful conceptual change.

Methods

Participants

Participants were recruited via an online advertisement through the university careers page. Participants were offered compensation of $20 (in retail vouchers) for their time. A total of 24 participants completed the study, all of whom were undergraduate students from The University of Melbourne studying in a range of areas. Participants were predominantly commerce or arts students. None of the students who completed the study reported having previously studied cosmology or physics and therefore had little to no prior scientific knowledge of the content of the learning task. The mean age of participants was 22.6 years. Eight participants were male and 16 were female. Ethics approval for this study was granted by the institutional Human Research Ethics Committee.

Materials

A mixed methods approach was used in this study. The measures of overall experience used in the study reported here are consistent with those we have used in previous studies and have been useful in gauging the global impressions participants had of the learning tasks they completed (see Lodge & Kennedy, 2015). Qualitative and quantitative questionnaires were supplemented by the use of video stimulated recall. Multiple measures were used to provide a deeper analysis of student progression through a discovery-based environment and offer suggestions as to when students were experiencing particular subjective states, such as confusion, during the task compared to their post-session reflections.

This study was conducted in a purpose-built computer laboratory on a 2012 model Apple iMac 27-inch computer running OSX ‘Mavericks’ operating system. The stellar lifecycle task (described below) was housed in a web-based interface provided by the *Smart Sparrow* platform run on Apple Safari web browser. Screen recording software (built in to Mac OSX operating system) was used to capture the sessions. All other materials (as described below) were provided on paper and were entered into Microsoft Excel and IBM SPSS software packages for analysis.

Stellar Lifecycle Learning Task

The task used as the lesson of this research was part of a larger task called *Habitable Worlds* (*HabWorlds*) created by Anbar and Horodyskyj in the *Smart Sparrow* platform and used at Arizona State University. HabWorlds is a course designed to help students learn about “the formation of stars, planets, Earth, life, intelligence, technological civilizations” (Anbar & Horodyskyj, n.d.). For the purpose of the current study, the section on stellar lifecycles was chosen as it provided a case of a commonly misconceived notion dealt with in a discovery-based learning environment. The Stellar Lifecycle task is not a ‘pure’ discovery-based environment but has some structure and scaffolding incorporated into it, as described below. Regardless, it serves the purpose of allowing us to explore how students experience environments where they have some flexibility in how they will progress through the task.

For this learning task, participants were asked to work through a series of screens that conform to a predict-observe-explain learning design (White & Gunstone, 1992). That is, participants first make a prediction (prediction screen) about what they believe is the relationship between star size and lifespan. Before moving on, they are also required to fill in a free text field explaining why they chose to make the prediction they did. From there, participants move onto a screen where they create virtual stars in a simulation space and observe what happens to the stars over time (observation screen 1). This space is displayed in figure 1.

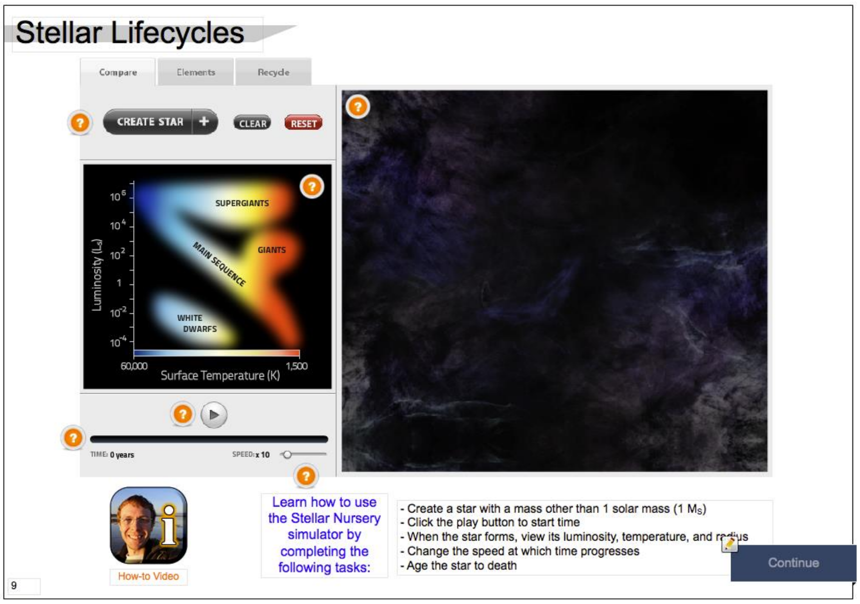

Figure 1. Observation screen for Stellar Lifecycle task

As can be seen in figure 1, there are numerous help options on the screen. There are also basic instructions on how to use the star simulator in the lower right hand portion of the screen. There is also an option for watching a 'how to' video in the lower left corner. A feature of the *Smart Sparrow* platform is that adaptive feedback can be given to students as they work through the tasks. The purpose of this screen is to allow students to create stars and see if their prediction holds. Students can run the simulator as many times as they wish before moving on. There is also the option of speeding up the simulator by manipulating the slide tool on the left side of the screen. Feedback and hints are given to participants if they are not using the stellar simulator effectively.

Upon completing the initial observation screen, participants then move to a more detailed simulation screen (observation screen 2), which has a very similar look and feel to the first observation screen. On this second observation screen, however, students are required to provide more detail in the simulator to create and age the stars. The main difference in the second observation screen is that participants are required to enter the mass and lifespan of the star. Again adaptive feedback was given to participants as they were completing the task, and as they were able to run through the simulator as many times as they wanted. Before moving on to the next screen, students were reminded of their initial prediction and asked if they still agreed with this or wanted to revise it.

Following the second observation screen, participants are simply asked to report what the correct relationship is between star size and lifespan by using a dropdown menu at the bottom of a screen (correct misconception screen). When they had completed this, participants were able to move on to the first of two explanation screens (explain screen 1 and 2). The first of these asked participants to estimate the shortest and longest lifespans of different classes of stars (according to standard classifications) in relation to mass. This estimation asks students to understand an extra layer of complexity to the content covered in the first few screens. While many stars fit within a main sequence where the relationship between surface temperature and luminosity is relatively consistent, some classes of star exist outside this main sequence. In the first explanation screen, participants are asked to incorporate this new information into what they have just learned. Examples of these classes of stars and their relationship to the main sequence are displayed on the left side of figure 1. White dwarf, giant and supergiant stars are all outside the main sequence so this detail is incorporated at this point in the task. The graph on the left side of figure 1 was also given to participants as a reference point on this first explanation screen. Again participants were provided with hints and feedback as they worked through the screen and were provided the option of attempting to enter the estimates as many times as they needed to.

A final screen was presented to participants to fully explain the lifespan of stars in relation to their mass. This screen included a video explanation, a reminder of their initial and corrected predictions and the correct estimates of the lifespans of different classes of stars. An overview of the learning design sequence of the task is presented in figure 2.

Figure 2. Overall learning design sequence for Stellar Lifecycle task

Video Stimulated Recall

The session was captured by screen recording software for a representative subset of 16 of the 24 participants. For practical and technical reasons, it was not possible to collect this data from all participants. These 16 participants were asked to complete a video stimulated recall session after they had completed their learning session. In this recall session, a research assistant played the screen capture of the learning session back to participants in its entirety. The video was stopped once every screen to allow participants to describe what they were doing, what they were experiencing, and what they were thinking at that point. The research assistant was trained in the use of prompting questions and took a handwritten transcription of the report of each of the participants about their experiences on each screen.

Survey

Participants completed survey instruments before and after completing the stellar lifecycle task to determine their overall impressions of their learning before and after the session. These instruments were based on similar instruments used in our previous studies (Lodge & Kennedy, 2015). Before completing the learning task, participants were asked to rate their confidence in their ability to learn about star lifecycles, how challenging they thought it would be to learn about star lifecycles and how much mental effort they felt they would need to invest in learning about star lifecycles effectively. These factors have been shown to be important considerations in the conceptual change process (Muller, Sharma & Reimann, 2008). Each of these ratings was made on a 0 – 10 scale with 0 representing “not at all” and 10 representing “very much so”. A second survey was used at the completion of the task. Participants again rated perceived confidence, challenge and mental effort on scales of 0 – 10 but were also asked to rate the degree to which they agreed that the task was interesting, enjoyable, confusing, frustrating and boring on Likert scales of 1 – 7 (with 1 representing “strongly disagree”, 4 representing “neutral” and 7, “strongly agree”). This survey also asked participants to provide demographic information.

Procedure

Participants were provided with information about the study and gave informed consent to participate as per the institutional ethics approval. Once completed, participants were given basic instructions about the sequence of the session. They were then given a survey to complete. After this, they were then permitted to complete the Stellar Lifecycle task at their own pace. At the conclusion of their learning session, participants were given the post-session survey to complete and then a subset of participants were asked to complete the video stimulated recall exercise. Once participants had completed all phases of the study, they were debriefed and given their compensation for completing the study.

Analytic coding of the responses to the prompts in the video stimulated recall session was conducted after the completion of the data collection. Thematic coding of the reported experiences of participants was conducted by one member of the research team and was then confirmed by a second member of the project team to ensure reliability of the coding process. The experience of participants on each screen was coded in such a way as to distil their experiences into a dominant theme (as per Merriam, 2009). Where participants reported experiencing several subjective states, the more prominent of these was chosen as the overarching theme being expressed at that point in the task. The resulting data therefore captures the prevailing experience of participants as they progressed through the task rather than a comprehensive survey of all of their subjective states at all points.

Results and discussion

Misconceptions

As expected, all participants except two incorrectly predicted the relationship between star size and lifespan on the prediction screen, indicating widespread misconceptions. Five participants reported explicitly that they took a random guess. The two participants who correctly predicted the relationship later reported they had guessed and that they did not know what the answer was. This supports the assumption that the relationship between star size and lifespan is a common misconception and, more to the point, the participants in this study showed this misconception. This observation also supports the notion that all participants in this study have started from a similar level of prior knowledge with the majority having an incorrect view of the concept in question.

Video stimulated recall

As reported, video stimulated recall was conducted with 16 of the 24 participants who completed the study. Summaries of the thematic analysis across participants for each screen are outlined below.

Prediction screen
As described above, the majority of participants made an incorrect assessment of the relationship between star size and lifespan. The main theme that emerged from the prediction screen was that the detail given on the screen confused participants, particularly the graph on luminosity that was included in the left side of figure 1. Nine of the participants reported that the graph confused them. The luminosity graph is only indirectly related to the star size-star lifespan relationship and the inclusion of the graph could be considered seductive details; details that are interesting but superfluous to the main intention of the lesson (as per Harp & Mayer, 1998). Whatever the cause, the observations from this screen highlight an important distinction between confusion that is caused by elements of the environment and confusion that is caused by the conceptual nature of the content. Future studies will need to be mindful of separating affective responses to elements of the environment and task from affective responses that are directly related to the conceptions at hand. This observation aligns with those we made previously in relation to tasks on pharmacodynamics and blood alcohol concentration (Lodge & Kennedy, 2015).

Observation screen 1
There was a mixed response to the information presented on this screen and several different themes emerged from participants' responses. While there was an overall theme of confusion and unhelpfulness in relation to elements of the screen, there also was an element of confidence and engagement in the reports, with seven participants reporting that they felt like they were making solid progress at this point. Two participants reported that they felt overloaded and confused as a result. Four others mentioned that they were explicitly confused at this stage. For example, one claimed "there is a lot going on in this page, I got confused". Two participants reported making the same errors multiple times by re-entering incorrect values into the simulator. There was more variance in the reports of participants as they worked through this screen than on the first. This again may be a reflection of different triggers (i.e. confusion triggered by the environment, the task or the concepts) and consequently of different affective trajectories.

Observation screen 2
The dominant theme to emerge from this screen was one of resolution. Six of the participants reported that they had developed a strategy of using specific intervals in the information they inputted into the star simulator. One stated: "I chose masses 2, 10 and 25 because it seemed the right range for low, medium and high". Despite this, there was still one participant who was confused about the luminosity by surface temperature graph. While there was still some sense that there was confusion among the group, there was a clear indication that the initial misconception was resolved as a result of the observations conducted in the observation screens.

Correct misconception screen
Participants spent very little time on this screen (14.8sec compared to the overall average screen time for all other screens of 153.9sec) and all managed to correctly select that smaller stars tend to have longer lifespans. The dominant theme to emerge from this screen was again one of resolution. Five participants explicitly stated that they were "very confident" in their knowledge by this stage. All participants managed to correct their misconception so the emergence of resolution as a theme is not surprising.

Explain screen 1
There was again some variability in the themes that emerged from participants' responses at this stage. While the dominant themes to emerge were again confusion, resolution and engagement, some sense of frustration was also expressed. One participant claimed that they "expected to be correct" while another claimed that they "didn't understand star classes and what they meant", further stating that they "found it really annoying". Nearly half of the group (seven) relied on the adaptive feedback provided on this screen to complete the task as they reported having difficulty filling out the approximate lifespans for each of the star classes so it is perhaps not surprising that frustration emerged as a theme here.

An extra layer of complexity was added at this point of the task (see explanation in "materials" above) as participants were asked to consider star classes that exist outside the main sequence. Ten of the 16 participants explicitly stated that they were confused by the numbers they needed to input into the screen (i.e. millions, billions and trillions of years). So while there had been one cycle of participants being confused and resolving the confusion and the misconception, this screen tended to plunge participants into a second round of confusion. The persistent experience of being confused may have led some to become frustrated, as predicted.

Explain screen 2
The main theme to emerge from the final screen in the task was again one of resolution. Eleven of the 16 participants reported that they were confident they understood the content of the task at this stage. Five further mentioned that they found the content presented in this final screen 'very interesting' and stated that they wanted to learn more. What this suggests is that participants were largely able to cycle through two rounds of confusion and resolution. Despite this, there was also some indication that some participants were frustrated and bored. Seven participants reported not paying attention to the video that provided further explanation of what the nature of the lifespan of stars is and this may have been a symptom of their boredom or frustration. Conversely, five participants felt that the video was important and watched it intently.

Individual pathways through the task
The thematic analysis conducted and described above was aggregated across all participants to provide a visual representation of the transitions between affective states participants experienced during the session. These are displayed in Figure 3.

As can be seen in the figure, participants reported various different affective experiences as they worked through the task. For example, one participant started off working through the task by being engaged but quickly became confused. She reported considering the information provided about luminosity when making a decision about which of the available options to choose. Moving along to the first observation screen, this same participant demonstrated that she was very much in a state of confusion. She reported:

> *"I got confused because data deletes when the star dies. 200 was not working, so I changed to fifteen. Didn't know what to do. Keep on trying. Noticing that smaller masses, life is longer. Try 300, still haven't realised. Tried "elements" and "recycle" because maybe I'm missing other options… not sure what I was doing. Didn't quite understand left graph and relationship to right."*

This participant then moved onto the second observation screen where she reported that she was developing an understanding of the correct relationship between star mass and lifecycle. She said “I was thinking I was pretty sure”, suggesting that she had managed to work through her confusion and had attempted to address whether or not her initial misconception was correct. When then reaching the correction of the misconception screen, this participant had little hesitation in changing the dropdown menu to reflect the correct star size to lifespan relationship. She then demonstrated an engaged approach to the explanation screen:

> *“Looking at numbers in the graph on the right. I change the numbers to match. I prefer 1 by 1 - my style. I use mouse to track - counting billions or trillion. I expected to be correct.”*

This description of the approach taken by the participant at this stage resembles the systematic approach taken by a proportion of students, as reported in Dalgarno et al. (2014). By this we mean that, similar to the observations in that study, this participant altered one variable at a time to determine “1 by 1” what the relationship between the variables is. This participant then completed the task by working through all the information in the second explanation screen. She reported that she found this part of the task interesting and wanted to learn more.

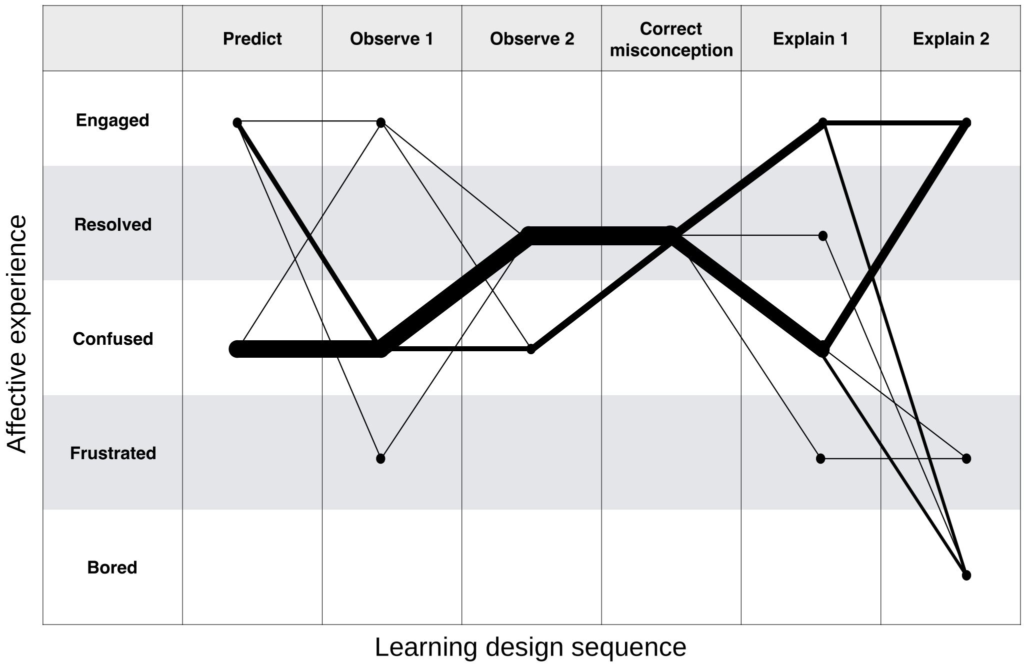

Figure 3. Affective transition diagram showing changes in participant experience through the stellar lifecycle task (heavier lines representing more participants following the path)

As can be seen in Figure 3, several other participants followed a similar pattern through the task (confused-resolved-engaged). However, a number of participants demonstrated a different pattern of responses. For example, two participants began in a state of confusion, resolved the confusion in the second observation screen and were initially engaged in the explanation of stellar lifecycles in explanation screen 1 but became bored by explanation screen 2. For example, one of the two stated that they “wanted to know more” suggesting they were engaged but then shifted to saying “I wasn’t so interested; wasn’t explained efficiently; lost interest”. Other participants were initially uncertain and confused, managed to resolve their confusion in the screen where they were asked to correct their misconception but became frustrated and disengaged by the time they reached the explanation screens. One participant claimed, “I got annoyed… it was too long… it went from very interactive to very ‘sit and watch’”.

These results suggest that this discovery-based learning environment, through its design, provided sufficient support and scaffolding to ensure that all participants effectively overcame their initial misconceptions or lack of knowledge about star lifespans and were able to achieve conceptual change. It seemed that, from the video stimulated recall and the recordings of students’ interactions within the environment, that the adaptive feedback played a role in keeping participants on track. In most cases, the feedback helped to correct information that was entered incorrectly into the screens. For example, the feedback given to participants on the first explanation screen helped them to set ranges for the classes of stars that were within the required range. The feedback given to participants took a form such as: "Your range of lifetimes for M class stars is incorrect. Run a star at 0.06M and one at 0.40M to determine the full range of ages for M class stars.” So while participants were relatively free to work through the task in their own way, the environment was not completely unstructured. Feedback was also given to students when they strayed too far from the "correct" path through the task. For example, participants were told if they had not completed important elements in the process such as leaving fields blank that should be filled in. These two mechanisms of feedback – adaptive pop up screens and hints about the sequencing of activities – are possibly the most important aspects of the design of this stellar lifecycle task from an affective perspective. That is, these built-in feedback mechanisms not only kept students focussed on the content and adaptive pathways within the learning task, they may also have warded off confusion and, importantly, persistent confusion that may lead to boredom and frustration.

When taken together the thematic analysis of students' affective responses from the VSR suggest that while students showed different learning actions and pathways though the task, and also different affective trajectories while traversing the material, there was some consistency in the "confusion to resolution" pathway. All roads lead to Rome.

Self report survey responses

While patterns were evident in the affective transitions participants made during the task, we also examined their overall impressions of the task. This provided an important contrast between participants shifting experiences during the task and their overall impressions.

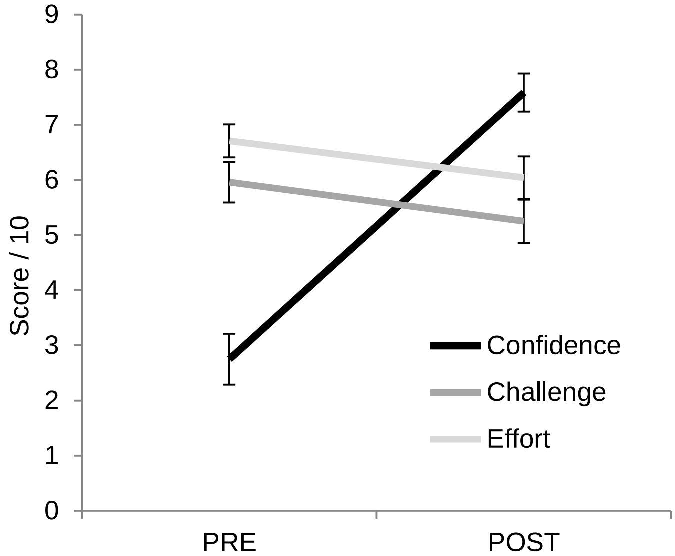

Figure 4. Mean (SE) pre and post-session ratings of confidence, challenge and effort

As can be seen in figure 4, participants reported relatively low levels of confidence in their ability to understand the material before the session (M = 2.75, SD = 2.3) compared to their assessment of how challenging the material would be (M = 6.0, SD = 1.8) and how much effort they would need to expend (M = 6.7, SD = 1.5). This relationship was reversed in the post task assessment. While there was a drop – but no significant difference – in participants' reported mental effort (F (1, 23) = 4.135, p = .054, partial eta squared = .15) and challenge (F (1, 23) = 2.33, p = .14, partial eta squared = .09) before and after the session, this was contrasted with participants reported confidence which significantly increased as a result of completing the session, (F (1, 23) = 75.27, p < .01, partial eta squared = .77).

Figure 5 shows participants ratings of their affective responses to the learning task after the session which indicate that they were more likely to report the session was interesting (M = 4.3, SD = 1.6) and enjoyable (M = 5.2, SD = 1.2) than confusing (M = 3.0, SD = 1.7), boring (M = 2.5, SD = 1.5) or frustrating (M = 2.4, SD = 1.5). While when asked to respond after the session participants predominantly indicated they enjoyed and found the task interesting, their experiences when prompted during the video stimulated recall suggested they spent a substantial amount of time being confused. Thus, there is evidence of a discord between measurements of affect completed through participants' global assessments and those completed through video stimulated recall.

Social desirability may in some way account for this; student participants may be more inclined, in a university context, to indicate to researchers they were more positive about the experiment and less confused or bored by it, in order to avoid any embarrassment or tension. However, leaving aside the potential of response bias due to social desirability, the findings also suggest that the specific affective transitions experienced by participants through the stellar lifecycles task (Figure 3) were not entirely reflected in the global assessment that participants made of their overall learning experiences (Figure 5). It is possible that retrospective, global ratings of affect inevitably ask students to commit to a "composite" self-assessment of their affect – students give a general estimate of their feelings about the entire learning session or task. And it is also possible that an individual's global assessment of affect could be different from assessments of their affective responses to specific components of the task. In effect, their assessment of the affective sum is different from their assessment of the affective parts.

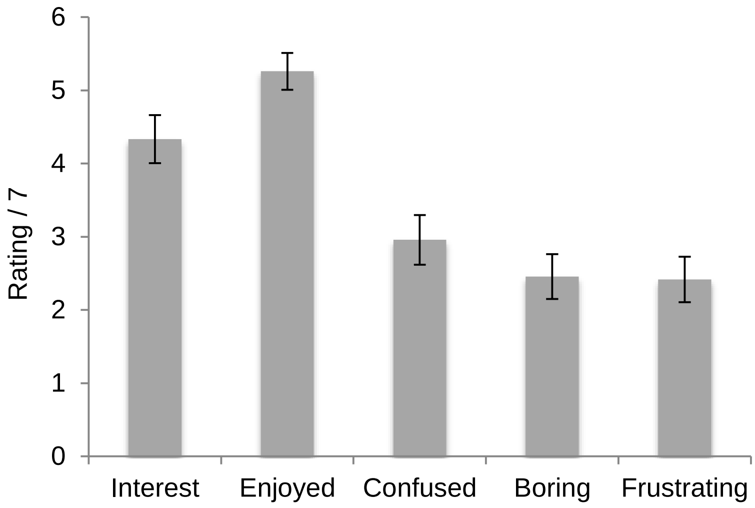

Figure 5. Mean (SE) affective ratings post-session

This potential raises important methodological questions about the investigation and measurement of students' affective responses to learning tasks. One interpretation of the findings from this study is that researchers may not be able to assess learning confusion by asking students at the conclusion of the task. That is, any assessment of affect needs to be closely tied with specific components and attributes of the task, as affect dynamically changes across open ended tasks and retrospective global measures are not able to capture this.

Researchers such as Rotgans and Schmidt (2014) and Ainley and Hidi (2002) provide useful rubrics that may act as the basis for further research in this area. These researchers, while investigating self-regulation and interest rather than confusion, considered how task-based changes in students' learning states impact on their learning strategies, processes and outcomes. Therefore, future research could fruitfully investigate the temporal, task-based nature of confusion and how the dynamics of confusion about the environment, about the task requirements and about the concepts interact to influence the student experience. This research would be useful in determining when students spend long periods of time being confused, which may lead to frustration and boredom (see Arguel, Lockyer, Lipp, Lodge & Kennedy, in press). Moreover, understanding, measuring and "seeing" specifically when students are confused during learning tasks, will be foundational to the provision of feedback to students both about their understanding of the content of the task and their approach to it, hopefully leading to the resolution of confusion and productive learning.

Author note: The authors wish to thank Prof Arial Anbar and Dr Lev Horodyskyj for providing us with the Habitable Worlds module for this research. We also wish to thank Sarah Brcan and Paula de Barba for their assistance in conducting this research and writing up this paper. The Australian Research Council provided funding for this research as part of a Special Research Initiative for the ARC-SRI Science of Learning Research Centre (project number SRI20300015).

References

Ainley, M. D., & Hidi, S. (2002). Dynamic measures for studying interest and learning. In I. P. R. P. M. L. Maehr (Ed.), *Advances in motivation and achievement: New directions in measures and methods*. (Vol. 12, pp. 43-76). Amsterdam: JAI.

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance

learning? *Journal of Educational Psychology*, *103*(1), 1–18. doi: 10.1037/a0021017
Anbar, A. & Horodyskyj, L. (n.d.). *Habitable Worlds (HabWorlds)*. Retrieved from www.habworlds.org
Arguel, A., Lockyer, L., Lipp, O., Lodge, J. M., & Kennedy, G. (in press). Inside out: ways of detecting learners' confusion for successful e-learning. *Journal of Educational Computing Research*
Brown, D. E. (1992). Using examples and analogies to remediate misconceptions in physics: Factors influencing conceptual change. *Journal of Research in Science Teaching*, *29*(1), 17-34.
Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31* (1): 21–32.
De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–201.
Dalgarno, B., Kennedy, G., & Bennett, S. (2014). The impact of students' exploration strategies on discovery learning using computer-based simulations. *Educational Media International*, *51*(4), 310–329. doi: 10.1080/09523987.2014.977009
D'Mello, S. K. & Graesser, A. C. (2014). Confusion. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education*. New York: Routledge.
D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, *29*, 153–170. doi: 10.1016/j.learninstruc.2012.05.003
Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414.
Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
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.
Kuhle, B. X., Barber, J. M., & Bristol, A. S. (2009). Predicting students' performance in introductory psychology from their psychology misconceptions. *Journal of Instructional Psychology, 36*, 119–124.
Lehman, B., D'Mello, S., & Graesser, A. (2012). Confusion and complex learning during interactions with computer learning environments. *The Internet and Higher Education*, *15*(3), 184–194. doi: 10.1016/j.iheduc.2012.01.002
Lodge, J. M. & Kennedy, G. (2015). Prior knowledge, confidence and understanding in interactive tutorials and simulations. In T. Reiners, B.R. von Konsky, D. Gibson, V. Chang, L. Irving, & K. Clarke (Eds.), *Globally connected, digitally enabled*. Proceedings ascilite 2015 in Perth (pp. FP:178-FP:188).
Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist*, *59*(1), 14.
Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation* (3rd ed.). San Francisco: Jossey-Bass.
Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, *92*(2), 278-296.
Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist*, 25(1), 19-33.
Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction*, *32*(C), 37–50. DOI: 10.1016/j.learninstruc.2014.01.002
Schwarzschild, M. (1958). *Structure and evolution of the stars*. New York, US: Princeton University Press
Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Winne, P. H. (2016). Identity and epistemic emotions during knowledge revision: A potential account for the backfire effect. *Discourse Processes, 53* (5-6), 339-370, DOI: 10.1080/0163853X.2015.1136507 .
White, R. T., & Gunstone, R. F. (1992). *Probing understanding*. London: Falmer Press.
Zimmerman, B. J., & Paulsen, A. S. (1995). Self- monitoring during collegiate studying: An invaluable tool for academic self- regulation. *New Directions for Teaching and Learning*, *63*, 13-27.

Please cite as: Kennedy, G. & Lodge, J. (2016). All roads lead to Rome: Tracking students' affect as they overcome misconceptions. In S. Barker, S. Dawson, A. Pardo, & C. Colvin (Eds.), *Show Me The Learning*. *Proceedings ASCILITE 2016 Adelaide* (pp. 318-328).

Note: All published papers are refereed, having undergone a double-blind peer-review process.