Check for updates

# Development of the Chemistry Mindset Instrument (CheMI) for use with introductory undergraduate chemistry students

Deborah L. Santos, [a] Jack Barbera [b] and Suazette R. Mooring *[a]

Chemistry education research has increasingly considered the role of affect when investigating chemistry learning environments over the past decade. Despite its popularity in educational spheres, mindset has been understudied from a chemistry-specific perspective. Mindset encompasses one's beliefs about the ability to change intelligence with effort and has been shown to be a domain-specific construct. For this reason, students' mindset would be most relevant in chemistry if it were measured as a chemistry-specific construct. To date, no instrument has been developed for use in chemistry learning contexts. Here we present evidence supporting the development process and final product of a mindset instrument designed specifically for undergraduate chemistry students. The Chemistry Mindset Instrument (CheMI) was developed through an iterative design process requiring multiple implementations and revisions. We analyze the psychometric properties of CheMI data from a sample of introductory (general and organic) chemistry students enrolled in lecture courses. We achieved good data-model fit *via* confirmatory factor analysis and high reliability for the newly developed items, indicating that the instrument functions well with the target population. Significant correlations were observed for chemistry mindset with students' self-efficacy, mastery goals, and course performance, providing external validity evidence for the construct measurement.

## Introduction

A variety of beliefs contribute to students' motivational behavior in chemistry courses, with some of these beliefs specific to the subject of chemistry. Certain beliefs may influence student outcomes more than others, mediated by motivational processes. Specifically, mindset beliefs are linked to student persistence in the presence of challenge and theoretically yield differential academic outcomes aligning with these beliefs (Molden and Dweck, 2006; Burnette *et al.*, 2013; Yeager and Dweck, 2020). Students are well aware that chemistry is known to be a challenging course and this reputation is perpetuated and confirmed by low course retention rates and lower grades relative to other courses (Harris *et al.*, 2020). As STEM educators seek to promote retention in STEM courses and persistence in STEM majors, understanding mindset is particularly relevant in these contexts. Mindset has been qualitatively shown to play a role in the formation of chemistry identity (Hosbein and Barbera, 2020), thus it is important in the retention of students in chemistry majors. There is also evidence to suggest that mindset can support increased STEM diversity through preferential benefits to students who would be more likely to experience stereotype threats in STEM courses (Aronson *et al.*, 2002; Good *et al.*, 2003; Fink *et al.*, 2018; Canning *et al.*, 2019).

To unravel the motivational relations responsible for differences in student outcomes, appropriate measures of each construct must be established. Several researchers have recently criticized the mindset meaning system (Burgoyne and Macnamara, 2021), the measurement quality associated with it (De Castella and Byrne, 2015; Lüftenegger and Chen, 2017; Limeri *et al.*, 2020a), or both (Dupeyrat and Mariné, 2005; Tempelaar *et al.*, 2015; van Aalderen-Smeets and van der Molen, 2018). Likewise, meta-analyses of the mindset literature have highlighted the inconsistencies of mindset as a predictor of achievement with undergraduate student populations (Costa and Faria, 2018; Sisk *et al.*, 2018). These inconsistent findings may point to inappropriate measurement of the mindset construct with the population of interest, indicating possible lack of validity. Additionally, work published by Santos *et al.* (2021) and Limeri *et al.* (2020a) found that undergraduate chemistry students interpret the terminology used in mindset instruments (*i.e.*, "intelligence") in a broad range of ways, which leads to potential response process concerns as some interpretations may have different implied malleabilities associated with them (*e.g.*, knowledge is inherently a grow-able quality). To avoid these varied interpretations and improve response fairness, less broadly defined wording can be used in mindset instrument items.

[a] *Georgia State University, Atlanta, Georgia, USA. E-mail: smooring@gsu.edu*
[b] *Portland State University, Portland, Oregon, USA*

Post-secondary students cannot be expected to hold the same views that primary and secondary students would have about a complex subject such as intelligence. It is likely that undergraduates hold a more multiplistic definition of intelligence as they increasingly realize that success can be achieved within a variety of different domains and using a variety of cognitive skills. This is supported by arguments that a domain-specific mindset measure is more appropriate at the undergraduate level within domain-specific contexts (Shively and Ryan, 2013; Scott and Ghinea, 2014; Little *et al.*, 2016; Gunderson *et al.*, 2017; Gorson and O'Rourke, 2019). Many domain-specific mindset studies in STEM have incorporated mindset measures that simply modify the item language from "intelligence" to terms such as "biology ability" (Dai and Cromley, 2014), "programming aptitude" (Scott and Ghinea, 2014), or "math intelligence" (Shively and Ryan, 2013). These types of modifications seek to improve the predictive power of mindset on STEM course performance or other outcomes but lack the qualitative justification necessary to suggest valid construct measurement. Buckley *et al.* (2019) demonstrated that students provide a broad range of definitions for intelligence within the technological domain through the qualitative exploration of characteristic behaviors of intelligent people in technology. These findings indicate that ideas about intelligence within a single domain can be complex for students to define. In addition to supporting the need for domain-specific mindset measures, these findings support infusing specified definitions of domain-specific intelligence within the instrument to yield more consistent interpretations. Based on these prior studies, it is reasonable to assume that chemistry intelligence is a unique and complex trait. Therefore, its meaning should be clarified for students when asked to report their beliefs, especially considering that many have a novice-level understanding of the field.

Another aspect of measuring mindset that has been questioned in recent years is the factor structure intended by typical mindset instruments (Lüftenegger and Chen, 2017). Mindset instruments are usually designed to measure two subfactors, entity and incremental theory beliefs (Dweck, 1999; Yeager and Dweck, 2020). Despite the two-factor design, mindset is often treated as a unidimensional measure when interpreting students' responses by using cutoff values or terciles to identify respondents with a fixed mindset (Hong *et al.*, 1999). Studies have shown inconsistent results in factor structure with some favoring a single-factor model and others favoring the intended two-factor structure (Gunderson *et al.*, 2017; Lüftenegger and Chen, 2017; van Aalderen-Smeets and van der Molen, 2018). As a further critique on the validity of measurement in many mindset studies, they often report quantitative results using mindset as a predictor variable yet do not provide evidence of valid usage of the instrument with the studied population as they tend to omit confirmatory factor structure analysis. Therefore, the validity questions raised here must be taken into account when measuring domain-specific mindset and have driven our development of a chemistry-specific mindset instrument.

## Theoretical framework

Mindset theory is a popularized term referring to students' implicit theories of intelligence. Students can hold entity or incremental theory beliefs about human traits such as intelligence, personality, and morality (Dweck *et al.*, 1995a, 1995b; Levy *et al.*, 1998). Entity theories are beliefs that the specified trait cannot change or is "fixed." Incremental theories are beliefs that a trait is malleable or can grow (Molden and Dweck, 2006). Entity theorists regarding intelligence or academic abilities generally place emphasis on innate ability and view effort as a sign of lacking necessary natural skills. Incremental theorists, on the other hand, view effort as a means by which to improve and thus obtain these skills (Dweck and Leggett, 1988).

Theoretically, this results in incremental theorists exhibiting "growth mindset behaviors" such as putting forth more effort and persisting toward success because they believe it to be more attainable relative to entity theorists. Alternatively, entity theorists are more likely to exhibit "fixed mindset behaviors" such as procrastinating, avoiding evaluation, and self-hindering to remove emphasis from their natural ability onto their willful actions (Molden and Dweck, 2006; Burnette *et al.*, 2013). These behaviors are self-protective responses to challenge that reflect ego threat, either as a result of interpreting challenge as a threat to their self-perceived value as intelligent individuals or confirming their negative self-perceptions. These relations suggest a link between mindset, self-efficacy, and achievement behaviors. Self-efficacy, the belief that one can achieve the desired outcome, has been shown to relate to mindset in several motivational analyses and thus is a useful variable to consider for demonstrating external validity (Komarraju and Nadler, 2013; Bedford, 2017; Lytle and Shin, 2020).

The originally proposed meaning system that students utilize based on their beliefs stated that achievement goals differ between growth and fixed mindset individuals (Dweck and Leggett, 1988). Achievement goals encompass two dimensions: mastery *versus* performance and approach *versus* avoidance (Elliot and McGregor, 2001). A student who sets mastery-approach goals is focused on increasing understanding of the content, while mastery-avoidance goals imply avoiding lack of understanding. Comparatively, performance-approach goals drive students toward achieving high grades, while performance-avoidance leads to avoiding poor grades. It has been proposed that growth mindset aligns with mastery-oriented goals and fixed mindset aligns with performance-oriented goals (Dweck and Leggett, 1988; Smiley *et al.*, 2016). Empirical support for the link between fixed mindset and performance orientation is weak, and rather most students report some degree of performance orientation (Leondari and Gialamas, 2002; Burnette *et al.*, 2013; Dinger and Dickhäuser, 2013; Karlen *et al.*, 2019). This trend may be due to the increased emphasis on high-stakes testing and grades-based assessment within modern education systems. Finally, as previously discussed, mindset has varying empirical predictive power on achievement measures such as grades, yet theoretically, growth mindset should lead to improved grades through adaptive behaviors (Hong *et al.*, 1999; Blackwell *et al.*, 2007). Achievement goals and course performance variables offer additional potential for demonstrating external validity of appropriate mindset measures.

## Goals of study

The work described in this report represents one part of a larger mixed-methods study investigating the effects of mindset

beliefs in chemistry on various outcome variables. This portion addresses the development of a chemistry-specific mindset measure and the validity of data from an introductory undergraduate chemistry student population. It is crucial that the measurement of this variable be understood prior to its use in future studies focused on drawing conclusions about the effects of mindset on course outcomes and other aspects of student affect. The specific goals of this study were to:

(1) Develop an instrument specific to mindset regarding chemistry intelligence intended for introductory undergraduate chemistry students.

(2) Determine the reliability and validity of measurements made with the developed instrument when used in the target population.

The two research goals were carried out by addressing the following research questions:

(1) How can item wording be modified to produce improved student response-process and construct measurement?

(2) How can the instrument's response-scale and dimensionality be modified to produce improved student response-process and construct measurement?

## Methods

### Participants

Different iterations of the surveys were administered during Fall 2020, Spring 2021, and Fall 2021 semesters to students enrolled in introductory chemistry courses (general and organic chemistry sequences) at a large southeastern US research-intensive university. The majority of course sections participated each semester and instructors agreed to provide a small amount of extra credit for students' completion of the surveys. Students who did not wish to participate in the research study were allowed to complete an alternative assignment or simply decline consent while completing the surveys and were credited the same number of points as those who consented to participate. Students were recruited for cognitive interviews during the Spring 2021 semester from the same courses that participated in the surveys. A compensation of $10 was provided to interview participants. This study was approved by the institutional review board prior to data collection.

Across semesters, the sample was consistently representative of the overall course demographics. The majority of students identified as female (69%), which is representative of the STEM course enrollment at the institution. Most students reported being in their third year (41%) and as a pre-professional or STEM major other than chemistry (90%). The samples were consistently representative of the racial and ethnic diversity at the university according to student reports (37% Black or African American, 28% Asian, 15% White, of non-Hispanic origin, 12% Hispanic, 7% other). Approximately half of the students (53%) reported eligibility for a Pell Grant, which can be used as an approximate indicator of lower socioeconomic status. And approximately one-third (34%) identified as first-generation college students.

The rates for student participation compared to enrollment in participating course sections are shown in Table 1 for all

**Table 1** Student survey participation totals and course response rates out of section enrollment from Fall 2020 through Fall 2021

| Semester | Timepoint | Total participants | Response rate (%) |
|---|---|---|---|
| Fall 2020 | Pretest | $N = 851$ | 45.4 |
| | Posttest | $N = 593$ | 30.5 |
| Spring 2021 | Pretest | $N = 595$ | 30.5 |
| | Posttest | $N = 513$ | 30.8 |
| Fall 2021 | Pretest | $N = 514$ | 46.5 |
| | Posttest | $N = 436$ | 67.6 |

surveys by semester. Participation rates in each instructors' section varied. A quality control procedure was used to flag careless responses through items that directed students to select a particular answer to verify they were paying attention to the content of the statements. After the removal of students who did not select the indicated response for quality control items, the remaining participants' data were analyzed.

### Data collection

**Survey details.** Surveys were administered online *via* Qualtrics© software through a link provided to students in their course management pages by their instructors. All pretest administrations were conducted within the first three weeks of the semester and posttest survey data was collected during the three weeks prior to the final exam. Surveys were administered over the course of three semesters, which included various iterations of the instrument. During the Fall 2020 posttest and Spring 2021 pretest administrations, two response-scale versions were directly compared by randomizing participants between two instrument versions (*i.e.*, Version 2 and Version 3) using logic within the software.

**Measures.** As the Chemistry Mindset Instrument (CheMI) was being developed, a variety of item wording versions and response-scales were trialed throughout the piloting and testing stages. The items and response-scales from each version can be found in Appendix A. The first version (Version 1) we implemented modeled Dweck's original 8-item instrument very closely (Dweck, 1999). Version 1 used both entity and incremental subscales with a 6-point Likert response-scale, but item wording was changed from "intelligence" to "chemistry intelligence." For example, the incremental item "I can always change my intelligence" was modified to "I can always change my chemistry intelligence." Version 2 contained 14 items (7 incremental and 7 entity) describing different aspects of chemistry intelligence, measured on a 6-point Likert scale. Version 3 used the same ability descriptions of chemistry intelligence as Version 2, but used a semantic differential response-scale with 6-points and was condensed to a unidimensional structure. The final modification in Version 4 was the use of a 10-point semantic differential scale. The final version of the CheMI (Version 4) contains 7 items, each incorporating a different aspect of chemistry intelligence as defined by students in an exploratory stage of this study (Santos *et al.*, 2021). For example, one CheMI item states:

*My ability to apply chemistry knowledge is something...*
*(I can't change at all) 1 2 3 4 5 6 7 8 9 10 (I can change a lot)*

Additional measures known to associate with mindset beliefs were also included in the surveys (Appendix B). Self-efficacy was measured using 6 out of the original 8 items from the self-efficacy subscale in the Motivated Strategies for Learning Questionnaire (Pintrich, 1991). Responses were reported using a 6-point Likert scale ranging from strongly disagree to strongly agree. Achievement Goals were measured using the $2 \times 2$ framework proposed by Elliot and McGregor (2001) and the 12 instrument items associated with it. The wording in these items was modified to reflect learning in a chemistry course by changing all references to "in this class" to "in chemistry." The four subscales in this instrument each contain 3 items ranked on a 6-point Likert scale ranging from strongly disagree to strongly agree. The four dimensions are called Mastery-Approach, Mastery-Avoidance, Performance-Approach, and Performance-Avoidance.

**Grades.** Instructors from each participating section provided a spreadsheet with grades for all assignments and assessments throughout the semester. This set of scores was used to compute formative and summative performance scores for each student. Formative performance scores incorporated assignments such as homework or writing tasks as well as any assessments during the course of learning such as quizzes or clicker questions. Summative performance scores incorporated all chapter exams and the final exam. To compare across sections with different instructors and grading schemes, all achievement scores were converted into $z$-scores so that the mean and standard deviation associated with that section was accounted for.

### Cognitive interviews

To investigate response-process validity between different versions of the chemistry mindset instrument as well as construct validity of the item wording, cognitive interviews were conducted with five students during Spring 2021. The students who participated in interviews represented the diversity of the overall sample when considering course level, year, gender, and racial and ethnic backgrounds. Interviews lasted approximately one hour and were conducted using a semi-structured protocol (Appendix C). Initially, students were presented with a sorting task that prompted them to create their own categories or groups using the chemistry intelligence definition terms from the instrument items (Fig. 1). Students were then asked to assign names to their categories and explain why they sorted the terms the way they did. Following the sorting task, students were asked to restate instrument items in their own words to express their interpreted meaning. They were then asked to select a response and explain their reasoning behind a response choice. After responding to multiple versions of the same item, students were asked to compare the response scales in terms of how they felt when
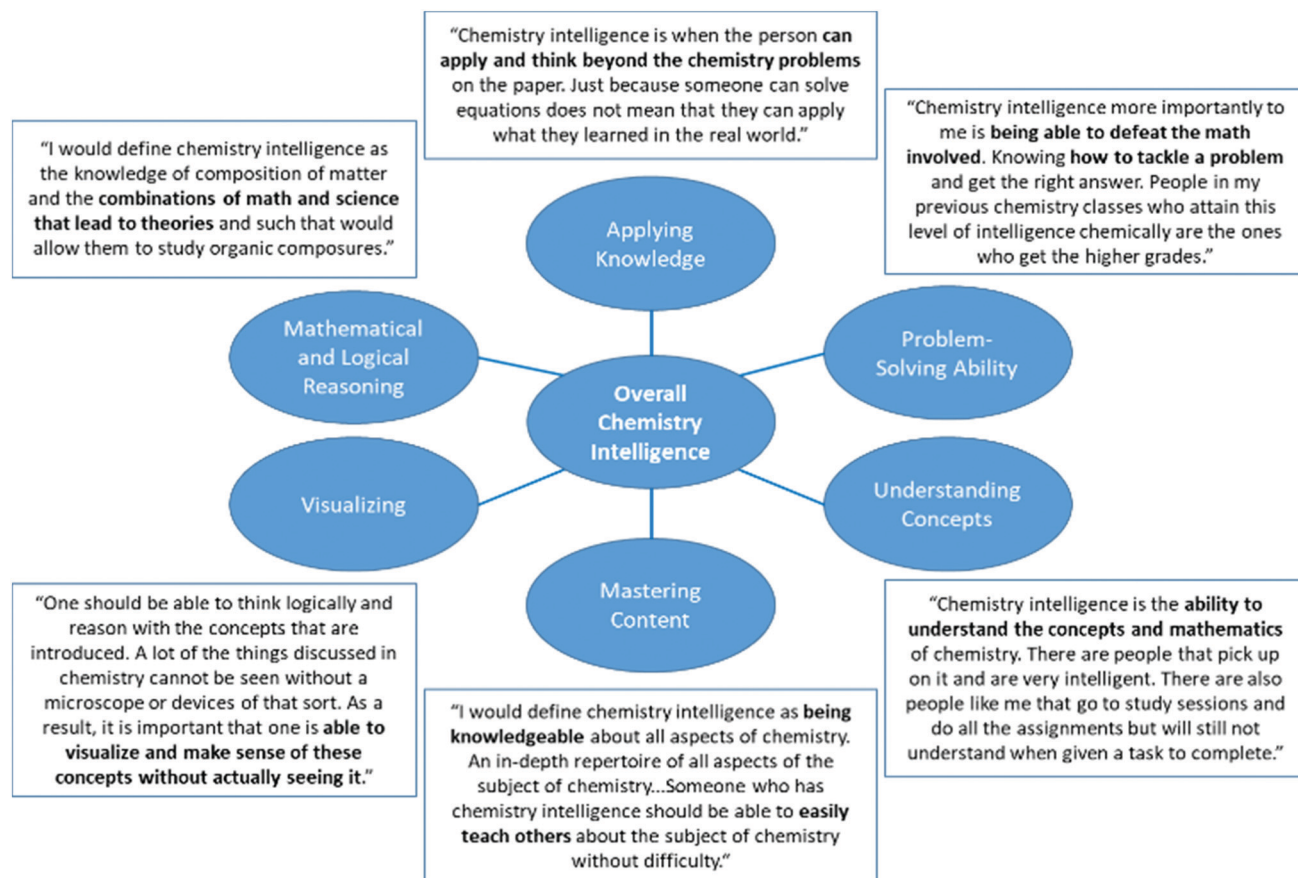


Fig. 1 Example student quotes that emphasize each aspect of chemistry intelligence included in instrument items.

selecting a response. All interviews were transcribed and analyzed for relevant perspectives on each type of response scale as well as interpretations of the items themselves.

## Data analysis

**Cognitive interview analysis.** All interviews were audio recorded and transcribed. The transcriptions were used to identify relevant comments on response scales and item wording. Any descriptions of feelings associated with a particular response scale or item wording were noted. Likewise, descriptions of wording or format influences on their decision to select a particular response and differences between responses when presented with a different scale were noted. Students' explanations regarding the meaning of a particular value were considered useful for determining their response processes across different item versions.

**Distribution normality and descriptive analysis.** All descriptive analyses were conducted using SPSS© version 28.0 software. As no significant differences have been observed in chemistry mindset mean scores between general and organic chemistry subsamples across instrument version administrations, all analyses were conducted on the full sample data. To analyze the response distribution for items and instrument versions, histograms were generated, along with computation of mean, standard deviation, skewness, and kurtosis values. When comparing separate versions, scale means were computed across all items. Versions with more central mean values were interpreted to show a reduction of social desirability and/or reduction in the ceiling effect of the response scale. Central tendency was expected due to claims that approximately a third of students in K-12 populations report a fixed mindset, which should theoretically yield a response below the middle of the scale (Hong *et al.*, 1999). Standard deviation was used to consider how students might allow some variability in beliefs regarding different items. More variation could indicate more careful and thoughtful responses to each individual item. Skewness and kurtosis values closer to zero were desired to show improvement of distribution normality.

**Factor structure.** Confirmatory factor analysis (CFA) was used to investigate the data-model fit of finalized items as indicators of a single unidimensional chemistry mindset construct. Mplus© version 8 (Muthén and Muthén, 2017) was used to run all CFA models using maximum likelihood estimation methods. Standardized factor loadings were expected to be greater than 0.7 to indicate a strong relationship between the item and the latent factor (Kline, 2015). Criteria suggested by Hu and Bentler (1999) were used to evaluate data-model fit.

**Reliability analysis.** The single-administration reliability of response measurement across the items in each version was also considered. Although Cronbach's alpha is most commonly reported, it assumes that factor loadings for all items are equal (McDonald, 1981; Dunn *et al.*, 2014). As this was not the case in CFA results for any instrument version, McDonald's omega provides a more appropriate estimate for single-administration reliability (McDonald, 2013). Interpretation of omega values is similar to alpha in that a value closer to 1 indicates more reliable measurements.

**External validity analysis.** To consider the validity of the instrument for detecting theoretically relevant relations between mindset and other variables (Dweck and Leggett, 1988; Hong *et al.*, 1999; Blackwell *et al.*, 2007), Pearson's correlation values were calculated between scale mean scores on mindset and the other measures collected. This allowed for correlation values to be computed with each of the four achievement goal dimensions, self-efficacy, and formative and summative achievement scores. The correlation values also indicate significant relationships at Bonferroni corrected $p$ levels. The size of the correlation was considered in relation to the strength of each relationship as indicated in the literature.

# Results and discussion

### Wording revisions: how can item wording be modified to produce improved student response-process and construct measurement?

**Wording changes.** During the Fall 2020 pretest survey, the first iteration (Version 1) of the CheMI was tested. This version used "chemistry intelligence" wording in the 8 items presented to students. This initial wording modification was introduced after a prior data collection using the original Dweck mindset instrument yielded a response distribution heavily skewed toward growth mindset in both the incremental and entity subscales. Along with the updated wording, open-ended questions regarding the definition of several terms, including chemistry intelligence, were posed to students. The in-depth qualitative analysis of student responses regarding these definitions has previously been published (Santos *et al.*, 2021). The results from the qualitative analysis were leveraged, during the development of Version 2, to substitute vague language ("chemistry intelligence") and exchange them for more self-explanatory terms students commonly use to describe it (for example, "ability to apply chemistry knowledge"). Fig. 1 presents the wording substitutions selected based on the previous study results (Santos *et al.*, 2021) along with quotes from open-response items highlighting the meaning of each term.

**Evidence from cognitive interviews.** During cognitive interviews, students read and explained their responses to multiple versions of instrument items to support and further inform development decisions. When asked to respond to the 4 entity belief statements from Version 1 (Appendix A), Abraham said, "When I got to the third one, I was like, 'Okay, this is just the same repeated question.'" He continued to say that he based his answers to the subsequent statements on his response to the first two. In comparison, when responding to items with the different cognitive abilities (Fig. 1) inserted as definitions or aspects of chemistry intelligence (see Versions 2, 3, and 4 in Appendix A), Abraham took longer to respond to several items and varied his response value, depending on the ability mentioned in the statement. Elena responded quite differently to each Version 1 statement because of differences in the meaning of descriptive words like "really" or how absolute some statements seemed compared to others. The other students said the

**746** | *Chem. Educ. Res. Pract.*, 2022, **23**, 742–757

This journal is © The Royal Society of Chemistry 2022

Version 1 Dweck-style items all meant the same thing as one another and thus responded the same across all items. When responding to the cognitive ability items (Versions 2, 3, and 4), some variation in their responses was observed due to differences in beliefs regarding their ability to improve various aspects of chemistry intelligence.

When presented with a range of cognitive abilities relevant to learning chemistry (Fig. 1), derived from prior results on definitions for "chemistry intelligence" (Santos *et al.*, 2021), students in interviews agreed that all were important factors to intelligence in chemistry. A sorting task was introduced to the students, instructing them to categorize the 7 chemistry intelligence terms in whatever way they believed they fit together. In the process of sorting these terms, Benjamin viewed them as abilities that develop sequentially while learning chemistry and that they all fall under "overall chemistry intelligence" as an umbrella term. During that same task, Camille commented that overall chemistry intelligence can change depending on improvements in the other six abilities. She stated that half of the abilities are less changeable and the other half she labeled as the "growth part of chemistry." Abraham said that the term "overall chemistry intelligence" related to all six of the other cognitive abilities listed. When reading a statement regarding the ability to change one's problem-solving ability in chemistry, Abraham responded by discussing the extent to which he believed chemistry intelligence can change. When asked why he brought up chemistry intelligence, he stated:

"Problem-solving ability connects a good amount with me to chemistry intelligence because if you have the ability to sort of comprehend hard problems, you have a good understanding of chemistry and you have a better chemistry intelligence than other students do. But not, it's not like all revolving around that. I guess it's like a certain aspect of your chemistry intelligence, which is a big, big thing. But I do think problem-solving is a good portion of chemistry intelligence."

To support the shift from the Version 1 term "chemistry intelligence" to various definitions (Versions 2, 3, and 4), it is helpful to compare how students responded to the first and final item wordings in cognitive interviews. Students commented about the repetitive nature of the original Dweck-style items (Version 1). This insight, combined with their reported tendency to simply select the same response for all items in a category, suggests that students do not feel the need to consider each item individually, but rather aim to respond consistently. This trend was not observed when asked to respond to items containing the various cognitive abilities as definitions of chemistry intelligence (Versions 2, 3, and 4). For these items, it was clear that students spent more time considering the nuances in their own abilities and beliefs regarding those abilities, thus leading to more variation in answer selection and care in representing their views about each item. These findings support the response-process validity associated with final item wording of the CheMI.

The construct of intelligence is a complex trait and can be defined in many ways. This complexity also applies when referring to intelligence within the discipline of chemistry. Equal emphasis across a broad range of cognitive skills deemed important for intellectual success in chemistry courses was selected for this instrument to provide a multiplistic view of intelligence within the measure. Interviews revealed that when presented with all of these cognitive aspects of chemistry intelligence (Fig. 1), students agreed that they are all important and fit within the umbrella of "chemistry intelligence." When asked what they believe the term "overall chemistry intelligence" to mean, they tended to respond that it meant all of the previously discussed aspects combined. These responses from students support the inclusion of each of these definitions within the construct of chemistry intelligence as presented in the instrument items. Additionally, cognitive interview findings that students understood the meaning of "overall chemistry intelligence" to refer to the other presented terms suggests that, within the context of the instrument, students are directed to interpret the broader term in light of all of the cognitive abilities referenced. These findings provide face validity evidence for student interpretations of the item wording as representing the same construct intended, chemistry intelligence. It can also be argued that any guesswork associated with interpreting a vague term such as "chemistry intelligence" is reduced within the context of the multiplistic definition as presented.

### Response scale modifications: how can the instrument's response-scale and dimensionality be modified to produce improved student response-process and construct measurement?

**Response-scale changes.** Version 2 of the CheMI contained 14 items with updated wording and retained the two-factor structure and Likert response-scale used in Version 1. This version was tested at the end of Fall 2020 in the posttest survey along with a randomly assigned comparison version. Version 3 incorporated a shift from a Likert scale to a semantic differential scale and was also piloted during the Fall 2020 posttest. The two versions were directly compared by random assignment of each student to one version or the other. Version 3 was created with the goal of removing the issue of entity *versus* incremental beliefs yielding inconsistent factor structure fit for the two-factor and one-factor models, as suggested by Lüftenegger and Chen (2017). A semantic differential scale assumes a unidimensional structure yet allows students the freedom to choose a particular viewpoint or any intermediate value on the sliding scale. In Version 3, the items were converted to a semantic differential scale format to allow students to choose a response that completes the statement to express their belief.

Substantial differences were observed between the Likert scale (Version 2) and semantic differential (Version 3) response distributions and were used to select the semantic differential scale for all future iterations. This decision was further supported through cognitive interview evidence. One last aspect was considered to improve the responses and measurement quality. A ceiling effect may be present for some students when using a 6-point scale, regardless of scale type. The fourth iteration (Version 4) of the instrument, therefore, contained a 10-point semantic differential scale. To test the efficacy of the expanded scale, Versions 3 and 4 were randomly assigned during the

survey administered at the beginning of Spring 2021. More detail into the evidence and rationale behind each of the modification decisions made are presented in the following sections.

**Evidence from cognitive interviews.** When prompted to compare response scales during cognitive interviews, students did not have a preference for either the Likert or semantic differential, but they tended to say there was more freedom to express how they felt when reading each semantic differential item. Abraham said that he views the two versions as saying the same thing, but in a different way. Benjamin also expressed that the statements had the same meaning despite different formats, but he provided a more extreme response to the Likert scale version. When comparing the two, Benjamin stated that a 5 out of 6 seemed equivalent to strongly disagree on the Likert scale, despite acknowledging that 6 out of 6 technically should be the equivalent value. He also commented that there were differences in meaning between clauses such as "can't change much" and "really can't change." In addition, he described the Likert scale as "more personal," making him feel more vulnerable in his response and more strongly about the statement. However, one student, Desiree, felt it was easier to relate to the Likert-style statements rather than the "arbitrary" numbering in the semantic differential version. Elena described that the semantic differential makes her feel like she has to "lean" one direction or the other, while the Likert scale is "just choosing from a list." She did not know if one version is better or worse than the other.

Students tended to prefer the 10-point scale to the 6-point scale because it was more familiar and provided more room for variation in their beliefs between items. For example, Abraham said, "Because there's more numbers and there's like more ways to put my feeling into the question. So, I feel like there's more numbers, like, I can better gauge how I feel about this certain thing. And then the other one, because I feel like, when I say 5 (out of 6), it's more of a vague answer than whenever I say 8 (out of 10)." Desiree also expressed that the 10-point scale is more familiar when thinking of the way people often rate things out of 10. When responding to a 10-point scale item, Elena said that she would choose 5, which aligned with the self-doubt she had expressed previously in the interview. However, when reading the same item on a 6-point scale, she stated that she would choose either a 4 or 5, which is much closer to a growth belief response. She also stated that the 10-point scale is more precise for her to be able to express her feeling on the statement, while the 6-point scale requires her to be more "decisive." Camille said that a higher value on the 6-point scale equates in her mind to a slightly smaller value on the 10-point scale (5.95 out of 6 is the same to her as 9.5 out of 10). She also said, "So, it's something that's…1 to 10 is easier to be visualized, at least in my mind, than a 1 through 6, even though, like, in the end, it's still the same. I believe it can change a lot." Benjamin commented on the reason for selecting a higher value on the 1 to 6 scale relative to the 1 to 10 scale, "I guess when it's like a smaller number range It feels like it's like, more severe as the numbers go lower." Commenting on the precision of each scale, he added,

"I think the smaller scale kind of feels a little more limiting, like it almost over summarizes maybe. As for the 10-point scale,
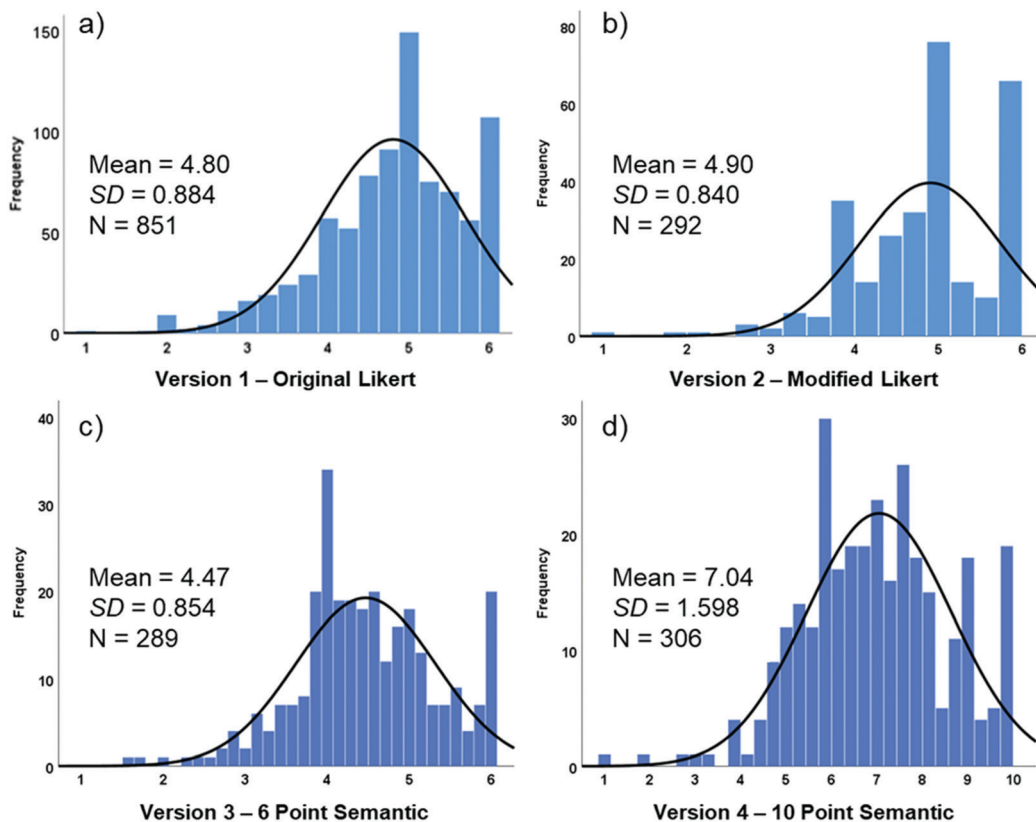
it might be able to be more specific. I mean, again, it's hard to say, because…I don't know what these – it's hard to say, like, what it even quantifies. I mean, because I'm just assuming, like, 6 and 10 are like infinity and then ones are nothings. Then it's like nothing to infinity."

The cognitive interviews provided useful evidence to support decisions related to students' response processes but did so most strongly for the transition from the 6- to 10-point semantic differential scale.

**Descriptive analyses.** Across the four piloted versions of the chemistry mindset instrument, mean distributions shifted in response to changes to the item wording and scale modifications. The full item-level descriptives across the four instrument versions are provided in Appendix A. The changes to the distributions can be seen in Fig. 2. The difference between sample distributions in the shift from Version 1 to Version 2 was not substantially improved. The following semester, Versions 2 and 3 were directly compared. As seen in Fig. 2c, the mean of the distribution for Version 3 is closer to the scale center relative to Versions 1 and 2, and more variation in item responses was observed. Lastly, Version 4 was noted as an improvement over Version 3 due to its increased central tendency and slightly larger standard deviation, suggesting that students may have responded more thoughtfully to individual items, increasing variability across items, as observed in cognitive interviews.

In addition to the improved central tendency of the mean across the instrument iterations, skewness and kurtosis both decreased with the modifications. These values are shown in Table 2. The smallest values for both skewness and kurtosis were observed in the final version (Version 4) of the instrument containing the 10-point semantic differential scale and the seven items with defined abilities. These values indicate a slight negative skew favouring growth mindset beliefs and slightly taller than a perfectly normal curve, but both fall well within the range of an acceptable normal distribution (Jones, 1969). The skew toward growth mindset has been observed consistently across versions and was most reduced in Version 4. This skew is most likely due to the social desirability of reporting growth mindset that has been noted in prior studies (Hong *et al.*, 1999; Santos *et al.*, 2021), and is likely impacted by the popularity of mindset instruction in K-12 learning contexts. The reduced skew observed with instrument modifications made here is likely due to a decreased in social desirability combined with the wording and response-scale modifications.

Response process validity. In comparing response-scale formats, students in cognitive interviews tended to select higher response values when using the Likert scale and said they had more freedom when using a semantic differential because it did not express a positive or negative viewpoint. These two findings support the claim that social desirability or acquiescence bias may play a role in response patterns for Likert versions of mindset items (Luftenegger and Chen, 2017). The less personal use of numbers without expressing a particular view, as seen in the semantic differential items, seems to influence student opinions to a lesser degree. Although the values themselves do not state their meaning (are "arbitrary,"

**748** | *Chem. Educ. Res. Pract.*, 2022, **23**, 742–757

This journal is © The Royal Society of Chemistry 2022

Fig. 2 Response distributions across four versions of the chemistry mindset instrument. (a) Version 1: Scale means for the Likert-scale 4-item incremental subscale (Fall 2020 pretest). (b) Version 2: Scale means for the Likert-scale 7-item incremental subscale (Fall 2020 posttest). (c) Version 3: Scale means for the bipolar 6-point semantic differential 7-items (Spring 2021 pretest). (d) Version 4: Scale means for the bipolar 10-point semantic differential 7-items (Spring 2021 pretest alternate).

Table 2 Skewness and kurtosis values for scale-mean response distributions across instrument versions

| Instrument version | Sample | Response scale/dimension | Skewness | Kurtosis |
|---|---|---|---|---|
| Version 1 | Fall 2020 pretest $N = 851$ | Likert/incremental | −0.837 | 0.761 |
| | | Likert/entity | 0.985 | 0.928 |
| Version 2 | Fall 2020 posttest [a]$N = 292$ | Likert/incremental | −0.760 | 1.391 |
| | | Likert/entity | 0.768 | 0.816 |
| Version 3 | Spring 2021 pretest [a]$N = 289$ | 6-Point semantic differential | −0.202 | 0.248 |
| Version 4 | Spring 2021 pretest [a]$N = 306$ | 10-Point semantic differential | −0.188 | 0.187 |

[a] These values are ~50% of those reported in Table 1 due to the random version assignment utilized during these administrations.

according to Desiree), they hold less value judgment and are left to the student to interpret. This provides evidence that Versions 3 and 4 reduce response-scale format influences on student responses, supporting response-process validity. These influences described in the cognitive interviews are likely a major cause of the skewed distributions observed in Versions 1 and 2.

Finally, consideration of response scale size was used to examine possible ceiling effects associated with limited value ranges. Upon initially responding to an item in cognitive interviews, students did not simply select the same scaled value between the 6-point and 10-point scale versions (*e.g.*, 5 out of 6 and 8 out of 10). They first selected their response, then

attempted to explain their response despite realizing it did not align with a direct conversion numerically. A less extreme value was selected when using the 10-point scale, indicating that having more scale options allowed them to feel more comfortable selecting a lower value. Students also said that a number lower than 5 on the 6-point scale seemed to be an "extreme" view to them, possibly indicating the effect of social desirability associations with a growth mindset leading to responses closer to the highest value. This finding aligns with the increased central tendency observed in Version 4 (10-point scale) relative to Version 3 (6-point scale), as shown in Fig. 2. Combining these results, we have evidence that a 6-point scale yields a ceiling effect for many student responses, and that this effect is

This journal is © The Royal Society of Chemistry 2022

*Chem. Educ. Res. Pract.*, 2022, **23**, 742–757 | **749**

reduced with the expanded 10-point scale. Reduction of a ceiling effect is beneficial in measurement to obtain better resolution of the distribution by shifting away from the scale edge and toward the center. Further support for the 10-point scale was provided in interview comments that a 1- to 10-point scale is more familiar and that it allowed students to be more precise in their responses, as evidence of response-process improvements. Considering all of this evidence led to the decision to retain a 10-point scale in the final CheMI version. The full final version (Version 4) of the CheMI is shown in Appendix A.

**Validity evidence for the CheMI Version 4**

**Internal structure and reliability.** Confirmatory factor analysis (CFA) was used to test that the unidimensional structure of the chemistry mindset construct aligned with all 7 items developed to measure it in Version 4. The data-model fit of the 7-item single factor model has been determined to be good across both data collections (Table 3).

Additionally, all items yielded high standardized factor loadings (all loadings $\geq 0.727$), indicating a strong relation between each and the overall latent construct. Fig. 3 presents

the CheMI Fall 2021 pretest data fit to a CFA model with standardized factor loadings. Similar CFA results were obtained in the posttest administration of the instrument in terms of the strengths of all factor loadings. To determine the single-administration reliability of responses across the 7 CheMI items, McDonald's omega ($\omega$) values were obtained for both the pre- and posttest survey administrations. Both time points yielded excellent reliability ($\omega_{\text{pretest}} = 0.929$, $\omega_{\text{posttest}} = 0.934$).

In addition to the cognitive interview evidence previously discussed that supports the construct alignment of all 7 items as measuring mindset beliefs about chemistry intelligence, CFA data-model fit and factor loadings corroborate the construct validity. All items strongly correspond to a unidimensional chemistry mindset construct, with no apparent subfactors. Additionally, students respond reliably across all items. As all 7 cognitive aspects appear to contribute to the overall construct according to multiple data sources, all 7 items were retained in the final version of the CheMI.

**Correlational analysis.** Correlations between data from the final iteration (Version 4) of the CheMI and other measures in the pre- and posttest administrations during Fall 2021 were

**Table 3** Data-model fit statistics across data collection timepoints using Version 4. Bolded values indicate results were good based on recommendations from Hu and Bentler (1999)

| Data collection | $N$ | $X^2$ (df, $p$) | RMSEA (confidence interval) | CFI | SRMR |
|---|---|---|---|---|---|
| Fall 2021 pretest | 514 | 29.34 (14, 0.009) | **0.046** (0.022 to 0.070) | **0.994** | **0.014** |
| Fall 2021 posttest | 435 | 28.90 (14, 0.011) | **0.049** (0.023 to 0.075) | **0.993** | **0.015** |
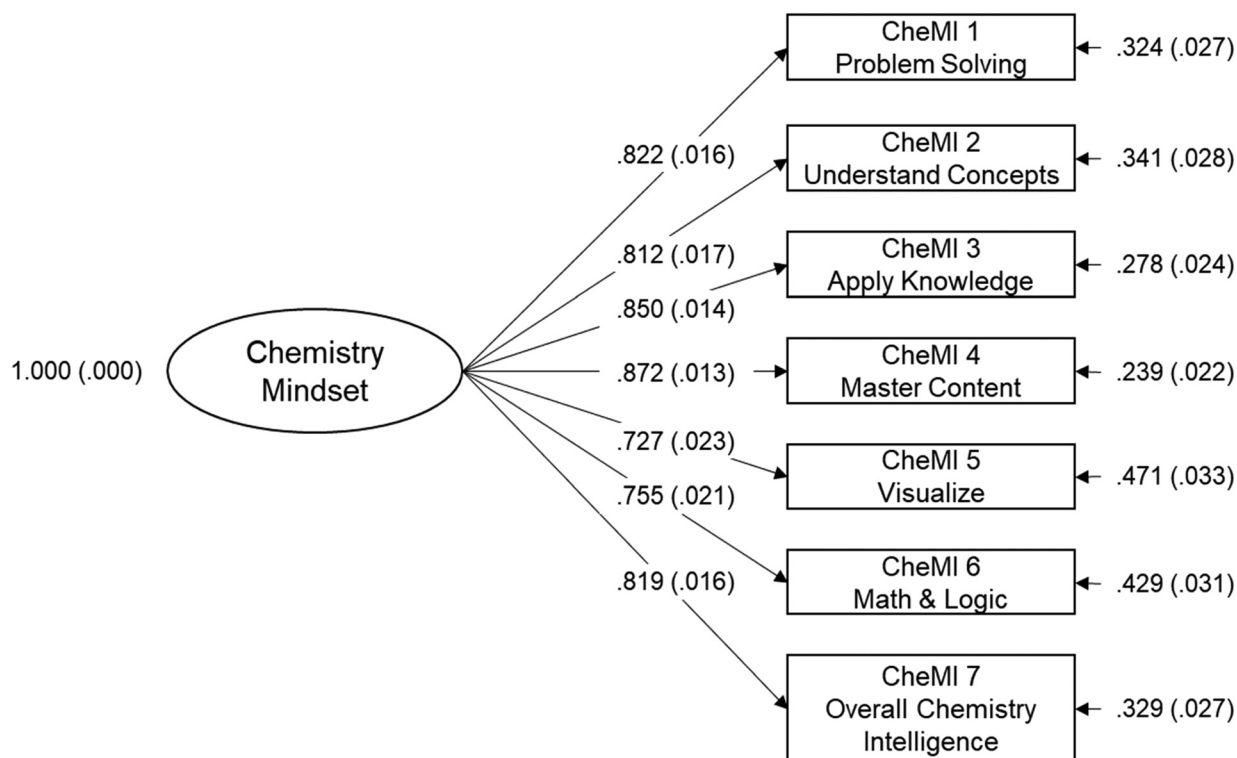


**Fig. 3** CFA model of Fall 2021 pretest Version 4 with standardized factor loadings. Item wordings can be found in Appendix A.

**750** | *Chem. Educ. Res. Pract.*, 2022, **23**, 742–757

This journal is © The Royal Society of Chemistry 2022

**Table 4** Pearson correlation values between chemistry mindset and other variables during Fall 2021

| Pretest | | | | | | Posttest Matched data, $N = 209$ |
|---|---|---|---|---|---|---|
| Variables | 1 | 2 | 3 | 4 | 5 | Chemistry mindset |
| 1 Chemistry mindset | 1 | | | | | **0.630**[b] |
| 2 Performance approach | 0.06 | 1 | | | | 0.167 |
| 3 Performance avoidance | 0.011 | 0.187[b] | 1 | | | −0.025 |
| 4 Mastery approach | **0.337**[b] | 0.124[a] | 0.147[b] | 1 | | 0.218[a] |
| 5 Mastery avoidance | **−0.139**[b] | −0.094 | 0.127[a] | 0.105 | 1 | **−0.226**[b] |
| 6 Self efficacy | **0.447**[b] | 0.134[a] | −0.007 | 0.409[b] | −0.247[b] | **0.475**[b] |
| **Course grades** | | | $N = 421$ | | | $N = 374$ |
| Formative scores | **0.168**[b] | 0.154[a] | −0.042 | 0.073 | −0.095 | **0.293**[b] |
| Summative scores | **0.228**[b] | 0.164[b] | −0.090 | 0.076 | −0.104 | **0.331**[b] |

[a] $p < 0.006$. [b] $p < 0.001$ for Bonferroni corrections with 8 variable correlations.

determined to provide evidence of external validity. These values are shown in Table 4. Bonferroni corrections were applied to all *p*-values due to the use of multiple correlations.

Self-efficacy yielded the largest external correlations with CheMI scores at both the pretest ($r = 0.447$, $p < 0.002$) and the posttest ($r = 0.475$, $p < 0.002$). This indicates that students with higher reported self-efficacy in their chemistry courses were more likely to report that they can improve aspects of their chemistry intelligence, aligning with findings from prior studies that these two constructs are positively related (Komarraju and Nadler, 2013; Bedford, 2017; Lytle and Shin, 2020). Likewise, mastery-approach goals were observed to correlate with both pre- ($r = 0.337$, $p < 0.002$) and posttest chemistry mindset ($r = 0.218$, $p < 0.002$), suggesting that students focused on mastery are more inclined toward beliefs associated with improving their chemistry intelligence (Dweck and Leggett, 1988). A negative correlation was observed for mastery-avoidance goals and chemistry mindset (pre: ($r = -0.139$, $p < 0.002$); post: $r = -0.226$, $p < 0.002$). Although the mastery-avoidance dimension of achievement goals was not a part of Dweck's original theoretical framework, the negative correlation with mindset can be expected because students' fears regarding their inability to learn the content align well with beliefs that chemistry intelligence cannot improve. No significant correlations were observed between chemistry mindset and either of the performance goal orientations. These results aligned well with previous findings that mindset more strongly relates to mastery-based achievement goals relative to performance-based goals (Leondari and Gialamas, 2002; Burnette *et al.*, 2013; Dinger and Dickhäuser, 2013; Karlen *et al.*, 2019). It should be noted here that a ceiling effect was observed for two of the four achievement goal dimensions: mastery-approach (found to significantly correlate with chemistry mindset) and performance-avoidance (not found to significantly correlate with chemistry mindset). Ceiling effects have also been observed in another study using the same achievement goals measure in chemistry (Lewis, 2018). The presence of these ceiling effects may limit the interpretability of the correlations observed; however, the expected relationships with chemistry mindset were observed, reducing this concern. Nevertheless, this

may indicate a need for an improved achievement goal orientation measure for chemistry-specific contexts.

The pre- and posttest mindset measures correlated significantly with both measures of course achievement, formative (pre: $r = 0.168$, $p < 0.005$; post: $r = 0.293$, $p < 0.005$) and summative scores (pre: $r = 0.228$, $p < 0.005$; post: $r = 0.331$, $p < 0.005$). As inconsistent results or small effect sizes have been observed in correlating mindset and achievement for undergraduate students across numerous studies (Sisk *et al.*, 2018; Costa and Faria, 2018), these findings were positive evidence of an improved mindset measure for chemistry contexts. Others have observed that including mediating variables between mindset and achievement yields significant predictive relationships (Macakova and Wood, 2020). However, lack of sensitivity of the mindset measure itself may further reduce direct predictive power, supporting that this instrument has increased sensitivity to differences in chemistry mindset beliefs.

Pearson correlations with pairwise deletion, theoretically relevant and significant values bolded for emphasis.

Examining the correlations between the CheMI measurements and other variables (Table 4) addresses one final consideration of validity, namely external validity. Convergent validity evidence is provided through the strength and sign of correlations with self-efficacy, mastery-based achievement goals, and course achievement according to the mindset meaning system (Dweck and Leggett, 1988; Dweck, 1999). Literature reports consistent alignment of mindset and mastery-based achievement goals (Burnette *et al.*, 2013) and a few studies have reported alignment with self-efficacy (Komarraju and Nadler, 2013; Bedford, 2017; Lytle and Shin, 2020). Theoretically, mindset beliefs in a domain should predict achievement (Hong *et al.*, 1999; Blackwell *et al.*, 2007), which is the primary incentive for conducting interventions. Yet, inconsistent findings in other studies with similar populations have brought these advantages into question for this academic stage (Sisk *et al.*, 2018). We argue that inconsistent findings may be a symptom of poor measurement quality for the target population's mindset construct, especially if the

domain-specificity of the construct has increased relevance for adult students. Divergent validity evidence was obtained by noting the near-absent correlations of mindset with performance-based achievement goals. Although a fixed mindset was originally found to relate to performance goals for young students (Dweck and Leggett, 1988), the increasing emphasis on performance as students progress toward high-stakes admissions processes is a possible cause for the lack of relationship between variables commonly reported in studies with secondary and tertiary students (Sisk et al., 2018).

Finally, several mindset-related studies in undergraduate STEM contexts have reported that domain-specific beliefs exhibit downward trajectories over time, indicating that students become more fixed in their beliefs (Dai and Cromley, 2014; Scott and Ghinea, 2014). However, the reported shifts in mindset are not large over a shorter time-scale such as one semester. This means that mindset beliefs at pre- and post-semester collection times should correlate strongly with one another. In our sample, pre- and post-chemistry mindset yielded the strongest correlation between variables (Table 4). This result should be interpreted with caution as many students were excluded from this correlation due to lack of participation at both timepoints ($N = 209$ for matched data). It does appear that some changes in students' beliefs did occur, as evidenced by the 0.630 correlation value. Students likely use their prior history with chemistry performance as evidence in the formation of their mindset beliefs at the beginning of the course. However, factors such as challenges, the classroom environment, and performance feedback in the current course may cause fluctuations and minor shifts in mindset throughout the semester. Although some students may have changed their views during the course as a response to their experiences and performance feedback (Limeri et al., 2020b), a single semester is a short time span for substantial changes in views.

## Conclusions

The Chemistry Mindset Instrument (CheMI) has been developed and shown to produce data that is valid and reliable according to multiple sources of evidence. The development and testing of this instrument was conducted with general and organic chemistry course populations. The instrument development process involved exploring literature suggestions for alternate response scales, open-ended responses to determine relevant definitions of chemistry intelligence for item wording modifications, cognitive interviews to determine response-process and face validity, repeated distribution and analysis of each iteration, and confirmatory factor analysis to verify appropriate fit of the data to the intended model for construct validity. Additionally, external validity evidence for CheMI data was provided through significant correlations with relevant variables such as mastery-approach goals, self-efficacy, and both summative and formative achievement scores. The CheMI was evaluated across two timepoints (i.e., early and late semester) to show that it yields data with reproducible psychometric properties and that reported

values correlate strongly with one another despite the passage of time. Students' post-semester chemistry mindset exhibited a stronger correlation with achievement variables, suggesting possible adjustment of beliefs during the semester to align with performance feedback in line with previous findings (Limeri et al., 2020b). The 7-item CheMI can be used to efficiently determine undergraduate students' chemistry mindset.

## Implications for research and teaching

Now that a chemistry-specific mindset measure has been developed and shown to produce valid and reliable data, it can be utilized to provide an understanding of the impact discipline-specific beliefs have on other relevant affective constructs. The length and simplicity of the CheMI is ideal for continued studies on the complex motivational pathways involved in student persistence and success in introductory college courses. Additionally, classroom interventions targeted at altering student mindset in chemistry or incorporating research-based teaching strategies can be monitored in terms of changes to chemistry-specific mindset beliefs. Students' native chemistry mindset belief trajectories in the absence of intervention can also be more adequately examined through longitudinal studies over the introductory course sequences. The CheMI can be useful to researchers, but also to chemistry instructors who wish to identify students who may be at risk for using maladaptive learning strategies as a function of their beliefs (Hong et al., 1999; Burnette et al., 2013). Once students are identified as having fixed mindset beliefs about chemistry, they can be supported with instruction about helpful study strategies, such as metacognitive strategies (Frey et al., 2020), mindset belief intervention assignments (Fink et al., 2018), and provided with positive messaging about investing effort and seeking assistance. Instructors may also wish to observe how changes to their teaching can impact student beliefs about learning chemistry. Studies have reported that instructor mindset can have a large impact on student outcomes and represent one of the factors that influence students' own mindset beliefs within that context (Canning et al., 2019; LaCosse et al., 2020; Muenks et al., 2020). Instructors may wish to observe how infusing mindset-related messaging impacts student beliefs about improving chemistry intelligence in their classes.

## Limitations

Correlational analyses were used as evidence of external validity in this study, but this technique does not allow for testing hypothesized causality or mediation effects of variables involved. Testing the mindset meaning system was not the focus of the work presented here, but rather verification that chemistry mindset measurements align with external variables as indicated in the literature. Future studies can examine the causal relationships among external variables such as motivational and behavioral measures using the CheMI through path modeling techniques. This can provide additional validity support by

752 | Chem. Educ. Res. Pract., 2022, 23, 742–757

This journal is © The Royal Society of Chemistry 2022

considering how data collected with this instrument fits within the hypothesized mindset meaning system. Additionally, this study only examined the CheMI's psychometric functioning with a student population from one institution, limiting the generalizability of the instrument's usage. To address this, chemistry mindset should be examined with students from other institutions and nationalities. During the development and evaluation of CheMI thus far, evidence has only been analyzed in aggregate. Therefore, future studies wishing to compare CheMI data across groups are encouraged to determine measurement invariance (Rocabado *et al.*, 2020). To date, this instrument has not been tested with students enrolled in courses other than general and organic chemistry, therefore validity evidence only applies to these introductory level courses. To expand its usage with additional populations, data collection and analysis with higher-level chemistry courses can be used to provide such validity evidence.

Variation in course participation rates was observed; however, the sample was representative of typical STEM course enrollment demographics at the institution across all categories. The different course sections were given different assignments and exams, thus raw average performance scores for formative and summative assessment may represent very different difficulty levels or assessment types. To mitigate this issue, $z$-scores for each course section were used to be more directly comparable relative to the performance distribution in each section. Additionally, all other measures were collected as self-report values, and thus may contain variation in interpretation and biases.

## Conflicts of interest

There are no conflicts to declare.

## Appendix A: item level descriptives across CheMI versions

| CheMI version | Mean | SD |
|---|---|---|
| **Version 1 (6-point Likert scale), Fall 2020 Pretest, $N$ = 851** | | |
| *Incremental items* | | |
| 1. No matter who I am, I can change my chemistry intelligence level. | 4.85 | 1.13 |
| 2. I can always change my chemistry intelligence. | 4.90 | 1.02 |
| 3. No matter how much chemistry intelligence I have, I can change it quite a bit. | 4.74 | 1.00 |
| 4. I can change my chemistry intelligence level significantly. | 4.73 | 1.12 |
| *Entity items* | | |
| 1. I have a certain amount of chemistry intelligence, and I really can't do much to change it | 2.22 | 1.08 |
| 2. My chemistry intelligence is something about me that I can't change very much. | 2.20 | 1.13 |
| 3. To be honest, I can't really change my chemistry intelligence. | 1.97 | 1.04 |
| 4. I can learn new things, but I cannot really change my level of chemistry intelligence. | 2.34 | 1.13 |
| **Version 2 (6-point Likert scale), Fall 2020 Postest, $N$ = 292 (randomly assigned 50%)** | | |
| *Incremental items* | | |
| 1. I can change my problem-solving ability in chemistry | 4.96 | 0.92 |
| 2. My ability to understand concepts in chemistry is something I can improve | 4.96 | 0.94 |
| 3. My ability to apply chemistry knowledge is something I can change | 4.96 | 0.90 |
| 4. My ability to master chemistry content is something I can improve | 4.98 | 0.94 |
| 5. I can improve my ability to visualize chemical structures and processes in chemistry | 4.76 | 1.08 |
| 6. My ability to use mathematical and logical reasoning in chemistry is something I can change | 4.77 | 1.06 |
| 7. My overall chemistry intelligence is something I can change | 4.93 | 0.99 |
| *Entity items* | | |
| 1. I can't really change my problem-solving ability in chemistry | 2.21 | 1.03 |
| 2. I can't change my ability to understand concepts in chemistry much | 2.27 | 1.05 |
| 3. My ability to apply chemistry knowledge is something I can't really improve | 2.27 | 1.08 |
| 4. My ability to master chemistry content is something I can't improve much | 2.14 | 1.05 |
| 5. I can't really improve my ability to visualize chemical structures and processes in chemistry | 2.37 | 1.09 |
| 6. My ability to use mathematical and logical reasoning in chemistry is something I can't change very much | 2.48 | 1.17 |
| 7. My overall chemistry intelligence is something I can't change | 2.25 | 1.16 |
| **Version 3 (6-point Semantic Differential), Spring 2021 Pretest, $N$ = 289 (randomly assigned 50%) (I can't change at all) 1 2 3 4 5 6 (I can change a lot)** | | |
| 1. My problem-solving ability in chemistry is something… | 4.49 | 1.04 |
| 2. My ability to understand concepts in chemistry is something… | 4.64 | 0.98 |
| 3. My ability to apply chemistry knowledge is something… | 4.44 | 1.06 |
| 4. My ability to master chemistry content is something… | 4.46 | 1.07 |
| 5. My ability to visualize chemical structures and processes is something… | 4.22 | 1.17 |
| 6. My ability to use mathematical and logical reasoning in chemistry is something… | 4.46 | 1.05 |
| 7. My overall chemistry intelligence is something… | 4.55 | 1.08 |

This journal is © The Royal Society of Chemistry 2022

*Chem. Educ. Res. Pract.*, 2022, **23**, 742–757 | **753**

Table (continued)

| CheMI version | Mean | SD |
|---|---|---|
| Version 4 (10-point Semantic Differential) Fall 2021 Posttest, $N$ = 436 (I can't change at all) 1 2 3 4 5 6 7 8 9 10 (I can change a lot) | | |
| 1. My problem-solving ability in chemistry is something… | 6.93 | 2.06 |
| 2. My ability to understand concepts in chemistry is something… | 7.08 | 2.05 |
| 3. My ability to apply chemistry knowledge is something… | 6.77 | 2.12 |
| 4. My ability to master chemistry content is something… | 6.83 | 2.28 |
| 5. My ability to visualize chemical structures and processes is something… | 6.48 | 2.12 |
| 6. My ability to use mathematical and logical reasoning in chemistry is something… | 6.91 | 2.03 |
| 7. My overall chemistry intelligence is something… | 7.00 | 2.15 |

# Appendix B: modified measures used in Fall 2021 pretest survey table

| Measure/construct | Original subscale | Modified subscale |
|---|---|---|
| [a]Achievement goal questionnaire/performance-approach | 1. It is important for me to do better in this class than other students<br>2. It is important for me to do well compared to others in this class<br>3. My goal in this class is to get a better grade than most of the other students | 1. It is important for me to do better *in chemistry* than other students<br>2. It is important for me to do well compared to others in *chemistry*<br>3. My goal in *chemistry* is to get a better grade than most of the other students |
| [a]Achievement goal Questionnaire/mastery-avoidance | 4. I worry that I may not learn all that I possibly could in this class<br>5. Sometimes I'm afraid that I may not understand the content of this class as thoroughly as I'd like<br>6. I am often concerned that I may not learn all that there is to learn in this class | 4. I worry that I may not learn all that I possibly could in *chemistry*<br>5. Sometimes I'm afraid that I may not understand *chemistry content* as thoroughly as I'd like<br>6. I am often concerned that I may not learn all that there is to learn in *chemistry* |
| [a]Achievement goal questionnaire/mastery-approach | 7. I want to learn as much as possible from this class<br>8. It is important for me to understand the content of this course as thoroughly as possible<br>9. I desire to completely master the material presented in this class | 7. I want to learn as much as possible from this *chemistry* class<br>8. It is important for me to understand *chemistry content* as thoroughly as possible<br>9. I desire to completely master the material presented in *chemistry* |
| [a]Achievement goal questionnaire/performance-Avoidance | 10. I just want to avoid doing poorly in this class<br>11. My goal in this class is to avoid performing poorly<br>12. My fear of performing poorly in this class is often what motivates me | 10. I just want to avoid doing poorly in *chemistry*<br>11. My goal in this *chemistry* class is to avoid performing poorly<br>12. My fear of performing poorly in *chemistry* is often what motivates me |
| [b]Motivated strategies for learning questionnaire/self-efficacy | 1. I believe I will receive an excellent grade in this class<br>2. I'm certain I can understand the most difficult material presented in the readings for this course<br>3. I'm confident I can understand the basic concepts taught in this course<br>4. I'm confident I can understand the most complex material presented by the instructor in this course<br>5. I'm confident I can do an excellent job on the assignments and tests in this course<br>6. I expect to do well in this class<br>7. I'm certain I can master the skills being taught in this class<br>8. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class. | 1. I believe I will receive an excellent grade in this class<br>2. I'm confident I can understand the basic concepts taught in this course<br>3. I'm confident I can do an excellent job on the assignments and tests in this course<br>4. I expect to do well in this class<br>5. I'm certain I can master the skills being taught in this class<br>6. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class |

[a] Elliot and McGregor (2001) A 2 × 2 achievement goal framework. [b] P. R. Pintrich (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).

# Appendix C: cognitive interview protocol used in Spring 2021

Spring 2021 Cognitive Interview Protocol: Chemistry Mindset Instrument

1. Student reads and signs the consent form before beginning the interview.

2. Researcher thanks student for participating and initiates with a few questions about their experience in the course.

3. Researcher instructs student to complete a series of several activities using the WebEx drawing tools and has the student project the activity documents on the screen as they work.

4. The student circles or crosses out items based on whether they correspond to their own beliefs about chemistry intelligence or behaviors in challenging chemistry scenarios.

5. After the student has responded to all items, the researcher will ask further questions and prompt for the next part of the activity, such as sorting remaining items into categories.

6. The researcher will ask the student questions about why they categorized items in this way.

7. The researcher will then show the student responses they gave in their survey earlier in the semester and ask about why they selected those answer choices (do they actually believe this or what other reason might they have chosen it?). Any discrepancies between responses on the survey and during the previous activities can be discussed.

8. The researcher will then ask the student to draw several graphs based on their own beliefs and discuss them in terms of comparing and contrasting their graphs/shapes.

9. The researcher will ask final open-ended questions to conclude the interview.

10. The student will acknowledge receipt of the gift card for participation.

Description of phases and questions students will be asked to respond to using think aloud:

Phase 1: Beginning questions to practice talking

– Personal challenge, effort, and engagement in chemistry – previous and present experience?

– Personal interest in chemistry, reason for taking it, and career goals?

– Meaning of natural ability?

– Interest and natural – can you have something naturally that doesn't interest you? Is interest natural or developed?

– Comes easily vs natural – are different things natural for different people?

Phase 2: Chemistry abilities sorting task (Fig. 4)

Provide instructions to the student that they should dictate aloud how they wish to sort the abilities into categories. They can create any number of categories as they wish.

– Ask students to label or name each category

– Ask for definitions of each term *in chemistry*

Phase 3: Chemistry mindset items

– Give Likert scale version first (Version 1). Ask student to respond to all aloud. Are they still reading each one?

– Ask how they would respond on a 10 point semantic scale (Version 4)

– Ask why they would choose that value, what does that number mean in words?

– Ask to compare 10 and 6 point scale (Version 3)

– Ask to compare Likert scale version (Version 2) – how would you respond and why? How does the format impact your answer or understanding of the item?

Items:

1. My problem solving ability in chemistry is something...

1 I can't change at all 2 3 4 5 6 I can change a lot

2. My ability to understand concepts in chemistry is something...

3. My ability to apply chemistry knowledge is something...

4. My ability to master chemistry content is something...

5. My ability to visualize chemical structures and processes is something...

6. My ability to use mathematical and logical reasoning in chemistry is something...

7. My overall chemistry intelligence is something...

Phase 4: Final questions

– Where do you think your chemistry intelligence comes from? Is this true for others?

– Do you believe your chemistry intelligence can change and what led you to believe that?
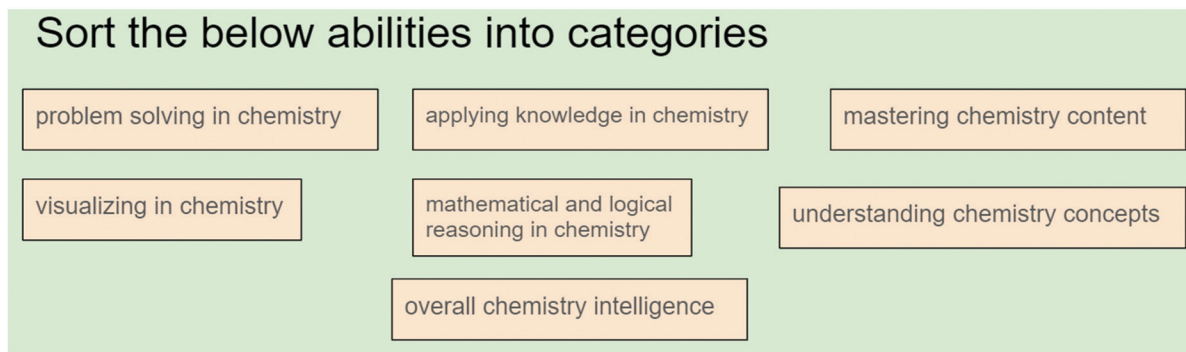


Fig. 4  Interview sorting task student view.

## Acknowledgements

## References

Aronson J., Fried C. B. and Good C., (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence, *J. Exp. Soc. Psychol.*, **38**(2), 113–125, DOI: 10.1006/jesp.2001.1491.

Bedford S., (2017), Growth mindset and motivation: A study into secondary school science learning, *Res. Pap. Educ.*, **32**(4), 424–443, DOI: 10.1080/02671522.2017.1318809.

Blackwell L. S., Trzesniewski K. H. and Dweck C. S., (2007), Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention, *Child Dev.*, **78**(1), 246–263, DOI: 10.1111/j.1467-8624.2007.00995.x.

Buckley J., O'Connor A., Seery N., Hyland T. and Canty D., (2019), Implicit theories of intelligence in STEM education: Perspectives through the lens of technology education students, *Int. J. Technol. Des. Educ.*, **29**(1), 75–106, DOI: 10.1007/s10798-017-9438-8.

Burgoyne A. P. and Macnamara B. N., (2021), Reconsidering the use of the mindset assessment profile in educational contexts, *J. Intell.*, **9**(3), 39.

Burnette J. L., O'Boyle E. H., VanEpps E. M., Pollack J. M. and Finkel E. J., (2013), Mind-sets matter: A meta-analytic review of implicit theories and self-regulation, *Psychol. Bull.*, **139**(3), 655–701, DOI: 10.1037/a0029531.

Canning E. A., Muenks K., Green D. J. and Murphy M. C., (2019), STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes, *Sci. Adv.*, **5**(2), eaau4734, DOI: 10.1126/sciadv.aau4734.

Costa A. and Faria L., (2018), Implicit theories of intelligence and academic achievement: A meta-analytic review, *Front. Psychol.*, **9**(829), DOI: 10.3389/fpsyg.2018.00829.

Dai T. and Cromley J. G., (2014), Changes in implicit theories of ability in biology and dropout from STEM majors: A latent growth curve approach, *Contemp. Educ. Psychol.*, **39**(3), 233–247, DOI: 10.1016/j.cedpsych.2014.06.003.

De Castella K. and Byrne D., (2015) My intelligence may be more malleable than yours: The revised implicit theories of intelligence (self-theory) scale is a better predictor of achievement, motivation, and student disengagement, *Eur. J. Psychol. Educ.*, **30**(3), 245–267.

Dinger F. C. and Dickhäuser O., (2013), Does implicit theory of intelligence cause achievement goals? Evidence from an experimental study, *Int. J. Educ. Res.*, **61**, 38–47, DOI: 10.1016/j.ijer.2013.03.008.

Dunn T. J., Baguley T. and Brunsden V., (2014), From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation, *Br. J. Psychol.*, **105**(3), 399–412.

Dupeyrat C. and Mariné C., (2005), Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults, *Contemp. Educ. Psychol.*, **30**(1), 43–59, DOI: 10.1016/j.cedpsych.2004.01.007.

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

Dweck C. S., Chiu C.-y and Hong Y.-y, (1995a), Implicit theories and their role in judgments and reactions: A word from two perspectives, *Psychol. Inquiry*, **6**(4), 267–285, DOI: 10.1207/s15327965pli0604_1.

Dweck C. S., Chiu C.-y and Hong Y.-y, (1995b), Implicit theories: Elaboration and extension of the model, *Psychol. Inquiry*, **6**(4), 322–333.

Dweck C. S. and Leggett E. L., (1988), A social-cognitive approach to motivation and personality, *Psychol. Rev.*, **95**(2), 256.

Elliot A. J. and McGregor H. A., (2001), A 2 × 2 achievement goal framework, *J. Person. Soc. Psychol.*, **80**(3), 501.

Fink A., Cahill M. J., McDaniel M. A., Hoffman A. and Frey R. F., (2018), Improving general chemistry performance through a growth mindset intervention: Selective effects on underrepresented minorities, *Chem. Educ. Res. Pract.*, **19**(3), 783–806, DOI: 10.1039/C7RP00244K.

Frey R. F., McDaniel M. A., Bunce D. M., Cahill M. J. and Perry M. D., (2020), Using students' concept-building tendencies to better characterize average-performing student learning and problem-solving approaches in general chemistry, *CBE—Life Sci. Educ.*, **19**(3), ar42.

Good C., Aronson J. and Inzlicht M., (2003), Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat, *J. Appl. Dev. Psychol.*, **24**(6), 645–662, DOI: 10.1016/j.appdev.2003.09.002.

Gorson J. and O'Rourke E., (2019), How do students talk about intelligence? An investigation of motivation, self-efficacy, and mindsets in computer science, Paper presented at the Proceedings of the 2019 ACM Conference on International Computing Education Research, Toronto ON, Canada, DOI: 10.1145/3291279.3339413.

Gunderson E. A., Hamdan N., Sorhagen N. S. and 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, *Dev. Psychol.*, **53**(6), 1188.

Harris R. B., Mack M. R., Bryant J., Theobald E. J. and Freeman S., (2020), Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone", *Sci. Adv.*, **6**(24), eaaz5687, DOI: 10.1126/sciadv.aaz5687.

Hong Y.-y, Chiu C.-y, Dweck C. S., Lin D. M.-S. and Wan W., (1999), Implicit theories, attributions, and coping: A meaning system approach, *J. Person. Soc. Psychol.*, **77**(3), 588.

Hosbein K. N. and Barbera J., (2020), Alignment of theoretically grounded constructs for the measurement of science and

**756** | *Chem. Educ. Res. Pract.*, 2022, **23**, 742–757

This journal is © The Royal Society of Chemistry 2022

chemistry identity, *Chem. Educ. Res. Pract.*, **21**(1), 371–386, DOI: **10.1039/C9RP00193J**.

Hu L. t and Bentler, P. M., (1999), Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Struct. Equ. Model.: Multidisciplinary J.*, **6**(1), 1–55.

Jones T. A., (1969) Skewness and kurtosis as criteria of normality in observed frequency distributions, *J. Sedimentary Res.*, **39**(4), 1622–1627.

Karlen Y., Suter F., Hirt C. and Maag Merki K., (2019), The role of implicit theories in students' grit, achievement goals, intrinsic and extrinsic motivation, and achievement in the context of a long-term challenging task, *Learn. Indiv. Diff.*, **74**, 101757, DOI: **10.1016/j.lindif.2019.101757**.

Kline R. B., (2015), *Principles and practice of structural equation modeling*, Guilford Publications.

Komarraju M. and Nadler D., (2013), Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learn. Indiv. Diff.*, **25**, 67–72, DOI: **10.1016/j.lindif.2013.01.005**.

LaCosse J., Murphy M. C., Garcia J. A. and Zirkel S., (2020), The role of STEM professors' mindset beliefs on students' anticipated psychological experiences and course interest, *J. Educ. Psychol*, DOI: **10.1037/edu0000620**.

Leondari A. and Gialamas V., (2002), Implicit theories, goal orientations, and perceived competence: Impact on students' achievement behavior, *Psychol. Sch.*, **39**(3), 279–291, DOI: **10.1002/pits.10035**.

Levy S. R., Stroessner S. J. and Dweck C. S., (1998), Stereotype formation and endorsement: The role of implicit theories, *J. Person. Soc. Psychol.*, **74**(6), 1421–1436, DOI: **10.1037/0022-3514.74.6.1421**.

Lewis S. E., (2018), Goal orientations of general chemistry students via the achievement goal framework, *Chem. Educ. Res. Pract.*, **19**(1), 199–212, DOI: **10.1039/C7RP00148G**.

Limeri L. B., Carter N. T., Choe J., Harper H. G., Martin H. R., Benton A. and Dolan E. L., (2020a), Growing a growth mindset: Characterizing how and why undergraduate students' mindsets change, *Int. J. STEM Educ.*, **7**(1), DOI: **10.1186/s40594-020-00227-2**.

Limeri L. B., Choe J., Harper H. G., Martin H. R., Benton A. and Dolan E. L., (2020b), Knowledge or abilities? How undergraduates define intelligence, *CBE—Life Sci. Educ.*, **19**(1), ar5.

Little A., Sawtelle V. and Humphrey B., (2016), Mindset in context: Developing new methodologies to study mindset in interview data, Paper presented at the Physics Education Research Conference Proceedings.

Lüftenegger M. and Chen J. A., (2017), Conceptual issues and assessment of implicit theories, *Z. Psychol.*, **225**(2), 99.

Lytle A. and Shin J. E., (2020), Incremental beliefs, STEM Efficacy and STEM interest among first-year undergraduate students, *J. Sci. Educ. Technol.*, 1–10.

Macakova V. and Wood C., (2020), The relationship between academic achievement, self-efficacy, implicit theories and basic psychological needs satisfaction among university

students, *Stud. Higher Educ.*, 1–11, DOI: **10.1080/03075079.2020.1739017**.

McDonald R. P., (1981), The dimensionality of tests and items, *Br. J. Math. Stat. Psychol.*, **34**(1), 100–117.

McDonald R. P., (2013), *Test theory: A unified treatment*, Psychology Press.

Molden D. C. and Dweck C. S., (2006), Finding ''meaning'' in psychology: A lay theories approach to self-regulation, social perception, and social development, *Am. Psych.*, **61**(3), 192.

Muenks K., Canning E. A., LaCosse J., Green D. J., Zirkel S., Garcia J. A. and Murphy M. C., (2020), Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability, engagement, and performance in class, *J. Exp. Psychol.: General*, **149**(11), 2119–2144, DOI: **10.1037/xge0000763**.

Muthén B. and Muthén L., (2017), *Mplus*, Chapman and Hall/CRC.

Pintrich P. R., (1991), *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*.

Rocabado G. A., Komperda R., Lewis J. E. and Barbera J., (2020), Addressing diversity and inclusion through group comparisons: A primer on measurement invariance testing, *Chem. Educ. Res. Pract.*, **21**(3), 969–988, DOI: **10.1039/D0RP00025F**.

Santos D. L., Gallo H., Barbera J. and Mooring S. R., (2021), Student perspectives on chemistry intelligence and their implications for measuring chemistry-specific mindset, *Chem. Educ. Res. Pract.*, **22**(4), 905–922.

Scott M. J. and Ghinea G., (2014), On the domain-specificity of mindsets: The relationship between aptitude beliefs and programming practice, *IEEE Trans. Educ.*, **57**(3), 169–174, DOI: **10.1109/TE.2013.2288700**.

Shively R. L. and Ryan C. S., (2013), Longitudinal changes in college math students' implicit theories of intelligence, *Soc. Psychol. Educ.*, **16**(2), 241–256.

Sisk V. F., Burgoyne A. P., Sun J., Butler J. L. and Macnamara B. N., (2018), To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses, *Psychol. Sci.*, **29**(4), 549–571, DOI: **10.1177/0956797617739704**.

Smiley P. A., Buttitta K. V., Chung S. Y., Dubon V. X. and Chang L. K., (2016), Mediation models of implicit theories and achievement goals predict planning and withdrawal after failure, *Motiv. Emot.*, **40**(6), 878–894, DOI: **10.1007/s11031-016-9575-5**.

Tempelaar D. T., Rienties B., Giesbers B. and Gijselaers W. H., (2015) The pivotal role of effort beliefs in mediating implicit theories of intelligence and achievement goals and academic motivations, *Soc. Psychol. Educ.*, **18**(1), 101–120.

van Aalderen-Smeets S. I. and van der Molen J. H. W., (2018), Modeling the relation between students' implicit beliefs about their abilities and their educational STEM choices, *Int. J. Technol. Des. Educ.*, **28**(1), 1–27.

Yeager D. S. and Dweck C. S., (2020), What can be learned from growth mindset controversies? *Am. Psychol.*, **75**(9), 1269–1284, DOI: **10.1037/amp0000794**.

This journal is © The Royal Society of Chemistry 2022

*Chem. Educ. Res. Pract.*, 2022, **23**, 742–757 | **757**