International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: http://www.tandfonline.com/loi/tsed20

A review and evaluation of the internal structure and consistency of the Approaches to Teaching Inventory

Jordan Harshman & Marilyne Stains

To cite this article: Jordan Harshman & Marilyne Stains (2017): A review and evaluation of the internal structure and consistency of the Approaches to Teaching Inventory, International Journal of Science Education, DOI: 10.1080/09500693.2017.1310411

To link to this article: http://dx.doi.org/10.1080/09500693.2017.1310411

Published online: 23 Apr 2017.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION, 2017
http://dx.doi.org/10.1080/09500693.2017.1310411

Check for updates

A review and evaluation of the internal structure and consistency of the Approaches to Teaching Inventory

Jordan Harshman[†] and Marilyne Stains[‡]

Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE, USA

ABSTRACT
This study presents a review from 39 studies that provide evidence for the structural validity and internal consistency of the Approaches to Teaching Inventory (ATI). In addition to this review, we evaluate many alternative factor structures on a sample of 267 first- and second-year chemistry faculty members participating in a professional development, a sample of instructors for which the ATI was originally designed. A total of 26 unique factor structures were evaluated. Through robust checking of assumptions, compilations of existing evidence, and new exploratory and confirmatory analyses, we found that there is greater evidence for the structural validity and internal consistency for the 22-item ATI than the 16-item ATI. Additionally, evidence supporting the original two-factor and four-factor structures proposed by the ATI authors (focusing on information transmission and conceptual change) were not reproducible and while alternative models were empirically viable, more theoretical justification is warranted. Recommendations for ATI use and general comments regarding best practices of reporting psychometrics in educational research contexts are discussed.

ARTICLE HISTORY
Received 22 September 2016
Accepted 20 March 2017

KEYWORDS
Psychometrics; validity; reliability; approaches to teaching; factor analysis

Introduction

Measuring beliefs about teaching

There has been a growing interest within the past decade to transform instructional practices in science, technology, engineering, and mathematics courses at the university level in the United States. Extensive research at both the secondary and postsecondary level has demonstrated that instructors' teaching practices are tightly related to their beliefs about teaching. The success of initiatives intended to train instructors in new instructional strategies can thus be demonstrated, in part, by a measured change in these instructors' beliefs about teaching. Teacher beliefs encompass several constructs that are often difficult to measure. However, some constructs, such as approaches to teaching, have been more extensively investigated than others. The popularity of investigating approaches to

CONTACT Marilyne Stains mstains2@unl.edu Department of Chemistry, University of Nebraska-Lincoln, 639 N. 12th Street, Lincoln, NE 68588, USA
[†]Dr Jordan Harshman is a first-year postdoctoral research in the Stains Research Group.
[‡]Dr Marilyne Stains is an associate professor of chemistry at the University of Nebraska-Lincoln.
Supplemental data for this article can be accessed here: http://dx.doi.org/10.1080/09500693.2017.1310411.

teaching is in part due to the relationship between these approaches and the actions of instructors, namely, the way instructors actually teach and how that impacts learning outcomes for students (Trigwell, Prosser, & Waterhouse, 1999). The three main arguments supplied by Martin, Prosser, Trigwell, Ramsden, and Benjamin (2002) nicely summarise this relationship. First, different teachers will have different intentions about what students should learn and will therefore vary in their approach to teaching. Second, what a student should learn is largely governed by teachers' perceptions of how students learn, which also impacts how they should be taught. Finally, if teachers intend for students to learn the material they are teaching, their actual pedagogy will relate closely with how they perceive is the best way to encourage learning. Essentially, understanding a teachers' approach to teaching is one crucial step in understanding the entire process of teaching and learning.

At the university level, relationships between conceptions of and approaches to teaching were the topic of research studies beginning around the 1990s (Kember & Kwan, 2000). Originally, five separate approaches to teaching were proposed (Prosser, Trigwell, & Taylor, 1994): teacher-focused with intention of transmitting information, teacher-focused with intention that students will acquire concepts, teacher–student interaction with intention that students acquire concepts, student-focused with intention that students develop their conceptions, and student-focused with intention that students change their conceptions. Since then, researchers have come a long way in discovering many aspects of these conceptions and approaches, but the theoretical dichotomy of teacher-centred versus student-centred is still largely proposed. The relationship between conceptions and approaches to teaching has been challenged previously (Kember & Kwan, 2000), but despite these cautions, is still used extensively (Sadler, 2012). This is likely because characterising the conceptions and approaches to teaching and learning has done such an efficient job of defining the role of teachers and students. These characterisations would dictate whether teachers saw learning as a process of information transfer or a more organic development of concepts, each implicitly defining different roles for the students as learners. While this dichotomous view would become widespread, it is not without disagreement. For example, Samuelowicz and Bain (2001) perceived an intermediate category defining student–teacher interactions. Åkerlind (2004) cautioned researchers against oversimplifying approaches to teaching and that there may exist many more aspects, a notion supported by Postareff and Lindblom-Ylänne (2008). A literature review (Sadler, 2009) contains more information for those interested in alternatives to the traditional teacher-centred/student-centred dichotomy. Being such an important concept, many have been interested in the measurement of teachers' conceptions and approaches to teaching, which has led to the development of a handful of quantitative instruments (Hativa & Birenbaum, 2000; Norton, Richardson, Hartley, Newstead, & Mayes, 2005; Trigwell & Prosser, 1996; Woolley, Benjamin, & Woolley, 2004). However, one instrument, the Approaches to Teaching Inventory (ATI) (Trigwell & Prosser, 1996), has emerged as a favoured tool among researchers to capture this abstract construct.

It is often the case in educational research that researchers will undergo the difficult task of designing and validating a survey, instrument, inventory, or tool for others to use. Once the initial instrument is developed and validated through one or more studies, the inventories are used in many other applications for many purposes. This pattern is observed with popular assessments such as the force concept inventory (Hestenes, Wells, &

Table 1. Studies examined.

ATI article	Citations	No. citations searched (time range)
Trigwell and Prosser (1996)	381	90 (1996–2005)
Trigwell et al. (1999)	1215	172 (1999–2005)
Trigwell and Prosser (2004)	401	401 (2004–2016)
Trigwell et al. (2005)	128	128 (2005–2016)
Prosser and Trigwell (2006)	153	153 (2006–2016)

Swackhamer, 1992, 2826 citations), the Science Teachers' Efficacy and Beliefs Instrument (Riggs & Enochs, 1990, 915 citations), and the subject of the current paper, the ATI (Trigwell & Prosser, 2004, see Table 1 for citation information). Given the importance of reproducibility in educational research, we were curious to know what evidence exists for the validity and reliability of this instrument after the initial development studies.

The Approaches to Teaching Inventory

Histories of the psychometric development of the ATI have been presented and critiqued previously (Meyer & Eley, 2006; Trigwell & Prosser, 2004), but to set the context of the present study, we provide a brief summary. Trigwell and Prosser conducted interviews with chemistry and physics university instructors to better understand their conceptions of teaching as reported in their 1994 article. It is from these findings that the authors first proposed the ATI in 1996. The ATI originated with 104 items, but was eventually cut down to 74, then to 49, 39, 22, 19, and ultimately 16 items (ATI16) for various reasons described throughout the main ATI development papers (Lindblom-Ylänne, Trigwell, Nevgi, & Ashwin, 2006; Trigwell & Prosser, 1996, 2004; Trigwell et al., 1999). It is important to note that the 22-item version reported in Trigwell & Prosser (1996) is not the same 22-item ATI used in our study, as this revised version came later in 2005 (ATI22, Trigwell, Prosser, & Ginns, 2005). The primary reason for developing a second version was twofold according to the authors (Trigwell et al., 2005). First, the authors believed that some items could benefit from rewording and one item (presumably item 1 on the ATI16) periodically caused confusion and should be removed. Second, they wished to 'extend the range of both constructs' (p. 353), but it does not appear as if the interpretation of the scales changes significantly from one version to the next.

The conceptual underpinnings and development of the ATI were critiqued by Meyer and Eley (2006). While we will not recapitulate all of these criticisms, we find two critiques of particular importance to the present study. First, the five conceptual approaches to teaching (Prosser et al., 1994) were developed based on interviews from just 24 physics and chemistry instructors, raising the concern of generalisability to other contexts and instructors. Meyer and Eley claimed that the target population of the ATI, originally chemistry and physics instructors, has been gradually generalised. In later uses of the ATI, the original authors and others use the ATI with all science instructors and all university instructors, which are beyond the population originally studied. Secondly, Meyer and Eley, echoing other researchers' concern (Åkerlind, 2004; Postareff & Lindblom-Ylänne, 2008; Sadler, 2009; Samuelowicz & Bain, 2001), criticised the notion that only two opposing approaches to teaching – information transfer/teacher focus (ITTF) and conceptual change/student-focus (CCSF) – exhausted the variety in conceptions of teaching, thereby casting doubt on this dualistic theory of teaching approaches. These three critiques seriously jeopardise the validity and reliability of results obtained from the administration of the ATI, and therefore warrant further investigation.

Our own quantitative investigation of the validity and reliability of the ATI confirmed these critiques. In particular, in our effort to gauge the effectiveness of the Cottrell Scholars Collaborative New Faculty Workshop (CSC-NFW) for university chemistry instructors in the United States (Baker et al., 2014; Stains, Pilarz, & Chakraverty, 2015), we administered the 22-item Approaches to Teaching Inventory-Revised (Trigwell et al., 2005) to several cohorts of chemistry faculty. As recommended (AERA, 2014), we first investigated the evidence for the validity and reliability of the instructors' responses to the ATI22 to ensure that the latent constructs defining approaches to teaching provide valid and reliable data in our sample. When the evidence failed to support the scales targeted by the authors, we explored prior studies that had used the ATI and found that at least *26 unique factor solutions* for the ATI had been presented and supported by various degrees of empirical evidence. Once we realised how many researchers had previously presented their own unique factor solutions for the ATI, we decided that it would be beneficial not only to our specific project, but also to the broader community using the ATI to compile the evidence available.

Purpose of the study

With growing popularity of both Trigwell and Prosser's ITTF/CCSF approaches to teaching theory and the use of the ATI16 (16-item version) and ATI22 (revised 22-item version) in research settings (Figure 2), this study seeks to provide a review of the psychometric evidence for the structural validity and internal consistency of the ATI16 and ATI22 (original wording of each instrument can be found in the supplemental information, Appendix A). Following this review, we present our own psychometric evidence upon administration of the ATI22 to a sample that we believe is an ideal representation of the population originally targeted by the ATI, which are faculty of one specific discipline. Finally, we make recommendations based on these findings for the proper usage of the ATI according to our results. Specifically, we focus on the following research questions:

- What is the evidence for the internal consistency and structural validity of the ATI16 and ATI22 across the many studies implementing these instruments (RQ1)?
- Which factor structure for the ATI22, whether novel or previously proposed, fits the responses of a sample of new chemistry faculty the best (RQ2)?

Being aware of the instrument's limitations was critical to the findings of our own study. From our perspective, if we deny the theory put forth by Trigwell and Prosser (that there are two main conceptions of teaching, CCSF and ITTF), then no degree of empirical evidence can properly make the ATI a useful instrument because its scales represent a defunct theory. Therefore, we focus primarily on the psychometric evidence for the structural validity and internal consistency of data produced by the ATI specifically under the assumption that the ATI is designed to produce a valid measure of approaches to teaching in accordance with Trigwell and Prosser's approaches to teaching theory. This also grounds our view that any novel factor structure proposed should either be in accordance with the ITTF/CCSF theoretical lens or explicitly state an alternative theoretical lens for which the factors fit.

Methods

Literature investigations

To find the majority of the studies that used and presented psychometric statistics for the ATI, we used Google Scholar to examine all studies citing one of five articles commonly cited for the use of the ATI (Table 1). We did not examine studies referencing the 1996 and 1999 papers after 2005 because we assumed that a majority of users of the ATI would then be citing one of the more recent articles. We also noted that the ATI was not widely used prior to 2005 (see Figure 2), serving as another indicator that the majority of ATI users would cite more recent articles. We assume that any professional studies using the ATI would give proper citation to one of these articles, but also performed general searches using terms such as 'approaches to teaching' and 'conceptions of teaching' to find additional studies. Finally, articles frequently referenced others who have used the ATI, which made them a good resource for finding additional studies. By no means are we assuming that we found *every* article, thesis, book, or other source using the ATI, but these methods allowed us to capture a majority of its uses.

The first author skimmed every study identified to determine if the authors collected responses to the ATI or not. It is worth noting that many articles cite the work by the ATI authors for the approaches to teaching framework, phenomenographic methods, and results from their studies and not for the actual ATI. This, along with a high degree of studies showing up multiple times (studies that cite two or more ATI articles), helps to explain how a couple of thousand articles only yielded a little over one hundred actual uses of the ATI. If an article only used the ATI and did not conduct or report any psychometric evidence, no additional data were collected beyond the full citation. Alternatively, if psychometric data were available, we collected all reported reliability statistics (coefficient alpha), factor structures tested and/or proposed, variance accounted for by exploratory factor analysis (EFA) or principal components analysis (PCA) computations, and common fit statistics used in confirmatory factor analysis (CFA, χ^2, χ^2/df, GFI, AGFI, CFI, TLI, RMSEA, SRMR, and Akaike Information Criterion). In addition, we collected information regarding the sample the ATI was administered to, including educational level and information about their discipline(s). Lastly, we also took a brief statement of the purpose of the article, which version of the ATI was used, including total item numbers and translations, and whether or not any items were modified beyond translations. All analysis was performed using R, version 3.3.0 (R Core Team, 2016), lavaan package 0.5–20 (Rosseel, 2012).

Cottrell Scholars Collaborative New Faculty Workshop

In addition to reviewing most of the evidence available in existing literature, we conducted our own study by administering the ATI22 to chemistry faculty participating in the CSC-NFW (Baker et al., 2014; Stains et al., 2015) as well as a control group. A total of 322 responses were originally gathered, but 37 were deleted because their total survey time was less than 5 min (the ATI was one of several inventories given and average survey time was around 15–20 min) and 18 were deleted because their responses resembled patterns on multiple scales (i.e. selecting the same alternative across all of the scales). More than half (140) of the final sample ($N = 267$) participated in the CSC-NFW during the data collection years, 2012–2015. These participants were either starting their first or

Table 2. Demographics of CSC-NFW chemistry faculty.

Year	N	$N_{CSC\text{-}NFW}$	$N_{Control}$	Male[a]	Female[a]
2012	110	25	85	67	40
2013	61	37	24	40	18
2014	61	43	18	44	16
2015	35	35	0	22	13
Total	267	140	127	173	87

[a]Some participants' sex not reported.

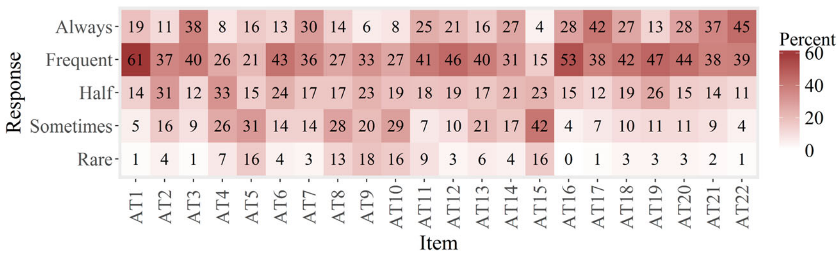

Figure 1. Distribution of the ATI22 responses in CSC-NFW data (percent shown in cells).

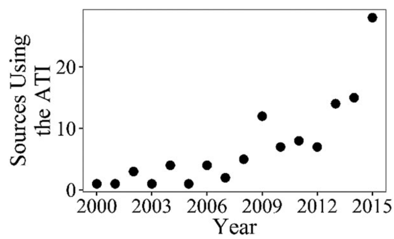

Figure 2. Use of the ATI by year.

second year as a chemistry faculty at a research-intensive university. The control group had a greater variety of experiences as faculty members than the CSC-NFW group but also worked at research-intensive universities. Because the instructors were all from chemistry departments, our sample closely matched those from the sample for which the ATI was developed (chemistry and physics faculty). Additional demographics are included in Table 2. A distribution of every item on the ATI22 is also provided in Figure 1. For some items, results show proclivity to select primarily three of the five response options while other items are more disperse.

Analysis

As an examination of internal consistency, we will be using coefficient alpha, which can be loosely thought of as an average of inter-item covariances (common cut-off $\alpha > .70$). For structural validity, we will be discussing results from EFA and CFA. We will provide a very brief overview of each here. EFA commonly involves the empirical and theoretical selection of a specific number of factors that are expected to exist in the data, where a factor is a group of items that measure a latent variable. Once the number of factors is chosen, a

Table 3. Cut-off values of CFA fit statistics.

Statistic	χ^2/df	CFI	GFI	AGFI	TLI	RMSEA	SRMR
'Best Fit'	<2	>0.95	>0.95	>0.95	>0.95	<0.05	<0.08
'Better Fit'	<3	>0.90	>0.90	>0.90	>0.90	<0.08	<0.10
'Poorer Fit'	>3	<0.90	<0.90	<0.90	<0.90	>0.08	>0.10

model needs to be selected for which to extract factor loadings. Factor loadings represent how well a particular item measures the factor it is loading on. Finally, a rotation method is used to rotate the axes of factor loadings in an attempt to maximise the number of items that load very poorly onto all but one factor for ease of interpretability. The purpose of this process is to end with a solution that indicates which items seem to measure the same latent variable. This is useful primarily as a means to explore natural groupings in items (hence the name exploratory factor analysis), but can also be used as evidence to confirm or refute a theoretical factor structure.

More commonly, researchers have turned to CFA to provide confirmatory evidence that a particular inventory measures one or more latent variables (hence the name confirmatory factor analysis). CFA belongs to a family of techniques called structural equation models, which model linear regressions simultaneously. The researcher defines a theoretical model and determines absolute fit (how well it predicts the observed data) and relative fit (compares how well it predicts the observed data versus how well a null model predicts the observed data) by considering a variety of fit statistics. Common fit statistics reported in CFA are the Chi square statistic divded by the degrees of freedom (χ^2/df), comparative fit index (CFI), goodness of fit index (GFI), adjusted goodness of fit index (AGFI), Tucker-Lewis fit index (TLI), root mean squared error of approximation (RMSEA), standardized root mean square residual (SRMR), and Bayesian information criterion (BIC). With several cut-offs proposed for each fit statistic to choose from, we primarily focus on those proposed by Hu and Bentler (1999), but include more and less stringent cut-offs listed in Table 3. CFA models can be altered by reassigning items to different factors, dropping items, and/or including error covariance terms. The expected change in fit statistics given this change can be found in modification indices, which can be used to help researchers 'fine-tune' their original CFA model. Lastly, all CFA results from present study are presented with Satorra–Bentler corrections to χ^2 due to evidence for non-normality.

Results

Demographics and factor structures identified in reviewed studies

We identified 101 studies in our literature search (see supplemental information, Appendix B), 39 of which presented at least some psychometric data. Unfortunately, this would indicate that over half of the studies reviewed used the ATI without reporting any evidence that the scales yielded valid and reliable data in their observed context. As alluded to previously, the ATI appears to be growing in popularity since 2007, as is displayed in Figure 2. This trend is likely due to the theoretical framework of approaches to teaching gaining popularity around 1999 (1215 citations) and the presentation of psychometric evidence of the ATI in 2006.

All of the results from here will be displayed only for the 39 studies that reported psychometric evidence. Sample sizes ranged from 20 to 2061 ($\mu = 278$, med = 177) and a total of 10,851 participants are reflected throughout the reviewed studies (data from the present study are included in the following results except where indicated otherwise). However, it is important to note that the studies administered the ATI to widely varying samples. A total of 31 studies administered the ATI to university faculty, accounting for 78.9% of the 10,851 participants reviewed. Of the eight studies that did not focus on university faculty, three studies targeted pre-service teachers and there was one study each that targeted graduate students, primary through secondary teachers, secondary teachers, university students, and secondary students. We coded the discipline of the participants as either precise (one or two subject areas, such as chemistry) or varied (multiple subject areas, not necessarily in science). Only 9 out of the 39 studies limited their investigation to one or two subject areas. The majority of the studies did not describe the disciplines of their participants or studied participants from many subject areas. The ATI16 was used in 21 of the studies, ATI22 was used in 16 of the studies, and 1 study was undeterminable. Finally, 15 studies translated the ATI into different languages (6 Spanish, 3 Chinese, 3 Dutch, 1 French, 1 German, and 1 Malaysian), potentially leading to changes in the ATI item meaning.

Excluding our own study, a total of 26 unique factor structure models were proposed either as a result of exploratory (EFA and PCA) or confirmatory (CFA) statistical procedures. This number does not include models that exclusively tested for unidimensionality (18 models), included items not native to the ATI (4 models), or had unknown factor structures (8 models). Models that tested unidimensionality were one-factor models that evaluated the evidence that all items belonging to that factor measure one construct. The reason that there were so many unique factor structures is because authors reported different numbers of factors as a result of EFA or PCA, different item memberships within factors evidenced by factor loadings, deletion of one or more items, and/or inclusion of error covariances within the context of CFA. While we will not discuss each of the 26 models proposed, we do focus on 9 of them (plus 2 original models we proposed) because they are either the most commonly used or show the largest degree of empirical support (see Table 4; the other 19 models are available in supplemental information, Appendix C). For example, Model 1 proposes that the ATI16 is composed of two scales, a CCSF and an ITTF scale. The CCSF factor in Model 1 is composed of items 3, 6, 8, 9, 14, 15, and 16 of the ATI16 and items 2, 4, 7, 10, 11, 12, and 13 load onto the ITTF factor. None of the models shown in Table 4 contained any error covariance terms.

The most common models evaluated in the literature were Model 1 (evaluated 12 times), Model 2 (6), Model 3 (3), and Model 4 (14). All other models were evaluated in only one or two studies. The original authors of the ATI have proposed Models 1, 2, and 4 as the suggested factor structure throughout their articles. As a result, it is important to note that Model 1 is the original two-factor structure composing a CCSF and ITTF scale for the ATI16. Model 3 is analogous to Model 1, except that it is based on the ATI22. Model 2 represents a four-factor structure, which breaks down the CCSF and ITTF factors into intention and strategy using the ATI16. Model 46 is analogous to Model 2 using the ATI22, but this structure was not officially proposed by the ATI authors.

Table 4. Models discussed throughout this manuscript.

Model	Structure: Factor	Structure: Items	ATI version
1	CCSF	3, 5, 6, 8, 9, 14, 15, 16	ATI16
	ITTF	1, 2, 4, 7, 10, 11, 12, 13	
2	CCSF (intention)	8, 15, 16	
	CCSF (strategy)	3, 6, 9, 14	
	ITTF (intention)	2, 4, 11, 13	
	ITTF (strategy)	7, 10, 12	
3	CCSF Scale		
	Conceptual Change Subscale	8, 6	
	Discussion among Students Subscale	6, 14	
	Discussion Between Students and Teacher Subscale	3, 9	
	ITTF Scale		
	Information Transmission Subscale	4, 7	
	Focus on Tests Subscale	2, 10, 11	
4	CCSF	3, 5, 7, 8, 13, 14, 15, 17, 18, 20, 21	ATI22
	ITTF	1, 2, 4, 6, 9, 10, 11, 12, 16, 19, 22	
5	CCSF (intention)	7, 14, 15, 17, 20, 21	
	CCSF (strategy)	3, 5, 8, 13, 18	
	ITTF (intention)	1, 2, 4, 10, 12	
	ITTF (strategy)	6, 9, 11, 16, 19, 22	
6	Making meaning	7, 17	
	Information transmission	1, 16	
	Focus on notes	4, 6	
	Focus on tests	11, 12	
	Discussions among students	5, 8	
7	Information transmission	4, 6, 16, 19	
	Preparing for tests	2, 9, 10, 11	
	Focused on students	17, 18, 20, 21	
8	ITTF	1, 2, 4, 6, 9, 10, 11, 12, 16, 19	
	Focused on students	7, 14, 15, 17, 18, 20, 21, 22	
	Focused on discussions	3, 5, 8, 13	
9	ITTF	1, 2, 4, 6, 9, 10, 12, 16, 19	
	Focused on students	7, 14, 17, 18, 20	
	Focused on discussions	3, 5, 8, 13, 15	

Evidence for the internal consistency of the ATI (RQ1, part 1)

To examine the first research question, we compared the evidence for internal consistency measured by coefficient alpha across the studies reviewed. To make a fair comparison, we first examined alphas only for the two-factor scale (Model 1 for the AT16, Model 3 for the ATI 22) and four-factor subscale (Model 2 for ATI16 and Model 46 for ATI22). These models were chosen because they reflect the original scales proposed by the authors of the ATI. We call this a fair comparison because the alpha is measured using exactly the same items in each study. This comparison can be found in Figure 3. Overall, 4 out of 24 studies reported an alpha of less than or equal to the commonly accepted, albeit arbitrary, .70 cut-off value (CCSF scale) and 7 of 24 fell below the cut-off for the ITTF scale. Only five studies reported alpha values for subscales, most of which were below the commonly accepted cut-off value. Finally, Figure 3 shows that the ATI22 generally outperforms the ATI16 on both the CCSF and ITTF scales in terms of internal consistency. In the supplemental information (Appendix D), we have reproduced Figure 3, this time including 16 models that are similar to, but not exactly the same as Models 1, 2, 3, and 5 (trends are generally the same).

At first glance, the data shown in Figure 3 might suggest that the ATI produces internally consistent data and the ATI22 is preferred based solely on internal consistency.

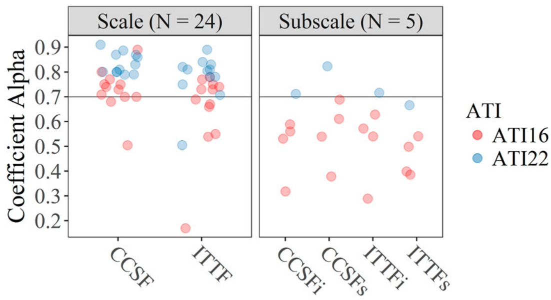

Figure 3. Coefficient alpha reported by studies using Models 1, 2, 3, and 5 (points are horizontally jittered for clarity).

However, we were recently made aware that coefficient alpha comes with assumptions that are rarely assessed (Barbera, 2016). In particular, the scale being measured is assumed to be unidimensional and tau-equivalent (Dunn, Baguley, & Brunsden, 2014; Graham, 2006; Trizano-Hermosilla & Alvarado, 2016). This means that each item in a factor measures one latent variable (unidimensionality) using the same scale and with the same precision (tau-equivalent). To exemplify this assumption, consider item 1 (‘In this subject students should focus their study on what I provide them’) and item 2 (‘It is important that this subject should be completely described in terms of specific objectives that relate to formal assessment items’), which are both assigned to the ITTF scale. The tau-equivalency assumption means that items 1 and 2 (and all other items on the ITTF scale) should (a) use the same Likert response options, (b) measure the ITTF construct with the same precision, and (c) measure this construct with different amounts of error. Assumption (a) is easily verified because they have the same Likert scale. Assumption (b) can be assessed empirically by fitting a unidimensional model to the data while constraining the factor loadings to a set value. Assumption (b) can also be assessed conceptually by asking, ‘would we expect Item 1 and Item 2 to load onto the ITTF factor in roughly equal ways?’ The last assumption, (c), is also assessed empirically by not constraining the errors and conceptually by asking, ‘would we expect the measurement error of these two items to be equal?’ If the model emerging from the more restrictive tau-equivalent assumption does not fit the data, then the less-restrictive congeneric model should be used. The congeneric model relaxes the assumption that all items load equally onto the factor, which seems to be a better assumption for the items on the ATI. Lastly, if the congeneric model is in fact a better fit of the data, coefficient alpha should not be used (Graham, 2006).

Only two studies presenting coefficient alpha also included evidence that support the unidimensionality of either the CCSF or ITTF scales, but neither of those assessed the assumption of tau-equivalency. This lack of checking assumptions could lead to global underestimations of the internal consistency of the ATI (internal consistency may be higher than reported) if a congeneric model fits better than the tau-equivalent one assumed by coefficient alpha (Graham, 2006). Because these assumptions were not widely assessed, we can only label our previous interpretations of Figure 3 as preliminary and provide further evidence with our own data set (see ‘Evidence of novel factor structure’ section). Lastly, coefficient alpha only measures one aspect of reliability, yielding a valuable, yet small component of the overall picture of validity and reliability of data produced by the ATI.

Evidence for the structural validity of the ATI (RQ1, part 2)

Focusing on the models originally proposed by the ATI authors (Models 1, 2, and 4), we examined the evidence for structural validity in the literature. Unfortunately, this evidence is sparse. Of the 12 studies reporting coefficient alpha for the two-factor Model 1, 5 verified this structure via EFA or PCA, 3 verified by CFA, and none of them did both. For Model 2, none of the six studies conducted EFA or PCA and three conducted CFA and for Model 3, two of the studies performed EFA or PCA and five performed CFA. Only four of the studies conducting EFA or PCA with Models 1, 2, or 3 included an estimate of variance accounted for by the model, which varied widely (41–80%). Several points are worth noting here. Firstly, we often could not tell whether EFA or PCA was run because of widespread confusion of the two that has been well documented (Meyer & Eley, 2006; Preacher & MacCallum, 2003). In the commonly used Statistical Package for the Social Sciences (SPSS), the default extraction is called 'Principal components', which leads to a PCA; the extraction method called 'Principal axis factoring' will lead to an EFA (UCLA: Statistical Consulting Group, n.d.). Second, several papers discussed a method whereby they discovered a larger number of factors (i.e. four-factor and five-factor solutions), did not report the results, and then forced a two-factor solution to verify the original models proposed by the ATI authors. This means that there are likely several more factor structures of the ATI that could support or refute other alternative structures.

For the CFA results presented in the literature, we compiled the fit statistics (when provided) for Models 1, 2, and 4, which are shown in Table 5. Two studies also referenced running CFA on Models 2 and 4, but did not report any fit statistics. Results presented from a variety of samples shows a generally poor fit for the two-factor and four-factor structures on the ATI16 and ATI22, although exceptions exist. This would suggest that the psychometric evidence for the factor structures originally proposed by the ATI authors is not consistently replicated. Our own results for the present study attest to a poor fit of these models as well (see 'Evidence of novel factor structure' section).

Evidence of previously identified factor structures (RQ2, part 1)

Having reviewed previously identified factor structures and evidence for internal consistency of the ATI16 and ATI22, we attempted to replicate these findings in our own data.

Table 5. CFA fit statistics reported for the three original models.

Model	χ^2	p	df	χ^2/df	CFI	GFI	AGFI	TLI	RMSEA	SRMR
1	172	<.001	62	2.8	0.88			0.85	0.08	
1	740	<.001	208	3.6	0.75	0.84	0.81		0.08	
1					0.87			0.84	0.06	0.06
2	113	<.001	72	1.6	0.96			0.96	0.05	
2	576	<.001	203	2.8	0.82	0.87	0.84		0.07	
4	477	<.001	208	2.3	0.85	0.80	0.78		0.08	
4					0.74			0.75	0.11	0.11
4					0.89	0.85	0.82		0.07	
4					0.95			0.94	0.06	0.08

We decided to first assess the internal consistency of the ATI22 according to the two-factor model (Model 4). Addressing the critiques mentioned earlier, we first assessed the unidimensionality of the CCSF and ITTF scales (Model 4) under the tau-equivalent and congeneric models, the results of which are shown in Table 6. From these results, neither the tau-equivalent nor the less-restrictive congeneric model fit the data well, suggesting that the unidimensionality of the factors outlined in Model 4 is suspect. We present our coefficient alpha and McDonald omega statistics for the sake of comparing our results with those reported previously, but argue that the two-factor model originally proposed by the ATI authors (Model 4) does not fit our data. To provide further evidence of the fit of Model 4 to our data, we ran a CFA of Model 3, also presented in Table 6. In consideration of the paucity of evidence for unidimensionality of each of the ITTF and CCSF scales, it is not surprising that Model 4 shows relatively poor fit to the data.

With multiple studies, including our own, failing to provide evidence for the existence of the originally proposed ITTF and CCSF factors, we identified three possible directions forward: (1) use modification indices from CFA and theory to add, remove, or reassign items to/from factors, (2) conduct EFA to investigate alternative factor solutions, or (3) choose not to use the ATI22 results in our analysis because we failed to provide evidence of the data's validity. The approach we actually took is a combination of the first two. Because we had already identified 26 unique factor solutions, we decided first to see if any of those solutions would fit our data. For models based on the ATI22, this was simple because we administered the ATI22. However, the ATI22 only contains 14 of the original ATI16 items. This meant that when using models based on the ATI16, we were unable to include the two items found on the ATI16 that are not included in the ATI22 (items 1 and 5). For example, Model 1 proposes that ATI16 item 1 belongs to the ITTF scale and item 5 belongs to the CCSF scale. We were unable to include these two items in our administration because we administered the ATI22, which does not contain these two items. Common fit statistics resulting from a CFA of each of the 26 models are shown graphically in Figure 4 and in tabular format in the supplemental information (Appendix E). Three models failed to converge (Models 3, 12, and 15) and are not included in Figure 4.

From the CFA results, it is clear that many of the models proposed previously do not fit our data well. Models based on the ATI22 performed better than models based using items found only on the ATI16. Empirically, Model 6 was the best model, passing the strictest cut-off criteria for all five fit statistics. However, Model 6 only contains 10 of the ATI22 items spread evenly across 5 factors. This is concerning because it has been argued that three observable variables are required to measure a latent construct (Hair, Black, Babin, & Anderson, 2010). We are in agreement with this assertion and therefore do not see Model 6 as a viable factor structure. The runner-up was Model 7, which just fell

Table 6. Results for unidimensionality under tau-equivalent and congeneric models.

Scale (model)	χ^2	p	df	χ^2/df	CFI	RMSEA (CI)	SRMR	BIC	α	ω
CCSF (tau)	167.3	<.001	54	3.1	0.829	0.089 (0.075–0.103)	0.106	8386	.83	–
CCSF (congeneric)	123.1	<.001	44	2.8	0.881	0.082 (0.067–0.098)	0.063	8394	–	.84
ITTF (tau)	147.5	<.001	54	2.7	0.797	0.081 (0.066–0.095)	0.101	7993	.80	–
ITTF (congeneric)	99.2	<.001	44	2.3	0.880	0.069 (0.052–0.086)	0.054	7994	–	.79
Model 4	456.8	<.001	208	2.2	0.802	0.067 (0.059–0.075)	0.074	16,390	–	–

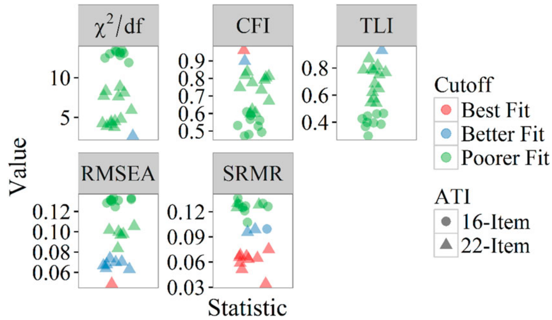

Figure 4. CFA fit statistics for 23 plausible models, colours represent cut-offs for statistic, points jittered for clarity.

short of a strict .95 cut-off for CFI and TLI (still above .90), but met the cut-offs for RMSEA and SRMR. Model 7 uses 12 of the ATI22 items spread evenly across 3 factors. Because no other models were able to achieve a CFI or TLI above .90, Model 7 was determined to have the best fit for our data.

Empirical fit is only one part of identifying a viable factor structure; the theoretical viability of the three factors had to also be considered. The first and last factors of Model 7, knowledge transmission and student-focused, are part of the ITTF/CCSF factors originally proposed, and therefore, are supported by the approaches to teaching theory. The examination preparation factor, however, is new and suggests that there is an approach to teaching that is centred around passing assessments. Chen and Brown (2016), the authors that proposed Model 7, did not provide any theoretical justification for the existence of an exam-focused approach to teaching, but three other studies found a similar factor related to assessment-focus (Goh, Wong, & Hamzah, 2014; Montenegro Maggio & González Ugaldeb, 2013; Stes, De Maeyer, & Van Petegem, 2010). Without a theoretical foundation to support the empirical results, Model 7 could be a spurious finding as opposed to evidence of a new perspective on approaches to teaching. Therefore, we decided to investigate our own factor structure.

Evidence of novel factor structure (RQ2, part 2)

In addition to evaluating models obtained from the literature, we also developed and evaluated models emerging from our own data. In fact, this chronologically occurred before we had looked at any other literature. An EFA was performed using promax rotation and maximum likelihood extraction. The ideal number of factors, 5, was determined via parallel analysis and accounted for 49.0% of the variance. By any standard, the five-factor solution was not a good fit, as two of the factors were composed of at most two items each. Assuming that additional factors would only exacerbate this issue of items failing to form coherent factors, we decided to examine the four-factor solution, but experienced a similar issue to the five-factor solution where one factor only contained two items (39.2% variance, see supplemental information (Appendix F) for four- and five-factor solutions). Lastly, we investigated the three-factor structure (34.6% variance) and have included the solution in Table 7.

Table 7. Three-factor solution.

Item	Factor 1 ITTF Scale	Factor 2 Focused on students	Factor 3 Focused on discussions
9	0.743		
11	0.665		
6	0.597		
4	0.591		
2	0.549		
10	0.518		
1	0.474		
12	0.390		
16	0.371		
19	0.351		
22		0.384	
13		0.354	0.415
18		0.459	
14		0.331	
21		0.572	
17		0.700	
7		0.644	
20		0.628	
15		0.464	0.376
8		0.413	0.574
3		0.356	0.457
5			0.718
Variance	0.138	0.268	0.346

It is interesting to note that in this solution, Factor 1 (ITTF Scale) is the ITTF Scale except without item 22. Factor 2 (focused on students scale) does not hold a clear interpretation, but seems to be more student-focused. Finally, Factor 3 (focused on discussions scale) pertains to holding discussion with and among students. This plausible factor structure, labelled as Model 8, was tested in a CFA, but showed poor fit ($\chi^2 = 421.1$, df = 206, CFI = 0.829, RMSEA = 0.063, SRMR = 0.075). A series of changes based on modification indices produced a better fitting model. Item 15 was reassigned to focused on discussions scale and item 22 was reassigned to ITTF Scale as suggested by modification indeces, improving the fit slightly ($\chi^2 = 401$, df = 206, CFI = 0.844, RMSEA = 0.060, SRMR = 0.071). Modification indices suggested a high correlation between items 9 and 11 (both assessment-focused) and items 14, 21, and 22 (which describe what students should do). Instead of building in error covariance terms, we eliminated items 11, 21, and 22 to reduce redundancy ($\chi^2 = 250.0$, df = 149, CFI = 0.900, RMSEA = 0.050, SRMR = 0.063). This final model, labelled Model 9, has three factors. The first mostly aligns with the original ITTF scale and another presents a scale similar to CCSF. However, a third factor is related to student discussions. Therefore, the model we discovered was similar to Model 7, in that there was a factor dedicated to information transfer and a factor related to student-focus, but a very different third factor. Similar scales that focused on discussions scale factors had been proposed in other studies using the ATI (Goh et al., 2014; Kreber, 2005; Stes et al., 2010). Therefore, Model 9 was somewhat supported empirically, but a theoretical justification is warranted before claims that such a factor structure exists.

Conclusions

Ignoring the limitations of coefficient alpha discussed previously, evidence was presented that the ATI22 has been demonstrated to produce more internally consistent results than the ATI16. Similarly, the ATI22 yielded better fit statistics in most models as compared to the ATI16 (Figure 4). Therefore, we conclude that if researchers choose to use the ATI, we recommend using the ATI22, as it yields better psychometric properties than the ATI16 (RQ1). As for determining the 'best' or 'true' factor structure of the ATI, there is not enough certainty for a global suggestion. First, it was demonstrated through multiple studies, the present one included, that evidence for the existence of the original two-factor and four-factor models proposed by the ATI authors (ITTF and CCSF subscales) is not reproducible. This should lead researchers to seriously question the use of the ATI as it was originally prescribed. Both the three-factor model proposed by Chen and Brown (2016, Model 7) and the three-factor solution we have proposed (Model 9) show preliminary evidence as viable structures, but without a deeper theoretical backing and reproducibility in other samples, no conclusions can be made with significant certainty regarding a factor structure for the ATI (RQ2). However, with multiple studies reporting evidence for the existence of factors outside of the traditional teacher-focused/student-focused dichotomy, we believe that future research should further investigate these factors. In particular, approaches to teaching that are focused on test preparation and focused on classroom discussions appeared in multiple studies and warrant further inquiry.

Lastly, it is worth noting that many models, including the two that produced sufficient evidence, contained a scale relating to the original ITTF scale proposed. Additionally, many other authors have proposed and provided evidence for the existence of something similar to an ITTF Scale. While this may provide evidence for the existence of an ITTF approach to teaching, we do not believe that such a construct is supported for three reasons: first, the evidence to support the ITTF Scale and those similar to it are not consistently reproduced. Several studies, including our own, sought alternative factor structure precisely because they were unable to reproduce the evidence for the ITTF (and CCSF) scales. Secondly, even when evidence exists for an ITTF Scale, which items actually measure it are highly variable. Our review found that authors believe that information transfer approaches to teaching could be measured by as few as 2 (Model 6) and as many as 11 (Model 4) items. With such a variety, it would be difficult to identify what this approach to teaching actually entails. Finally, as we have reviewed in the introduction, several authors have theoretical disagreements with the notion that there only exist two approaches to teaching. As a result of these three points, we are left to severely question both the notion of only two types of approaches to teaching and the existing tools used to measure these constructs.

Discussion

The ATI has had a remarkable impact on education research and will undoubtedly continue to be administered to instructors. Our experiences throughout this study have served as a humble reminder of how difficult it is to accurately measure latent constructs in education research and how important it is to examine the psychometric evidence of instruments repeatedly. In this light, we strongly discourage researchers from doing exactly as we did initially. After finding that the original scales proposed by the ATI authors did not fit our data, we were so quick to discover our own factor structure to fit the data that we neglected to consult the literature. But science does not occur in a vacuum and our results would have been incomplete without a thorough analysis of previously proposed structures. Therefore, we as a research community should aim to investigate what others have done with instruments prior to administering them ourselves.

Based on the results, we believe that the original authors of the ATI have done an exemplary job in introducing the first generation of measurement of approaches to teaching. Considering that the ATI was introduced over 20 years ago, its continued use is a testament to the impact it has had on research involving teachers' beliefs and more specifically, approaches to teaching. However our field has advanced our understanding of the complex subject of approaches to teaching and we have advanced our standards for the evidence that is necessary to deem an instrument valid in measuring a targeted concept. With these advancements, we look forward to researchers who will incorporate supplementary and/or alternative theoretical perspectives or to those who maintain the teacher-centered/student-centered framework and find novel means of measuring it.

Additionally, this study raises an interesting question in the topic of instrument development: what happens to instruments after they are released to the research community? Originally, the ATI authors provided evidence of the two-factor and four-factor structure, but how do we as a community track the further evidence in support of or refuting this structure? A wide diversity of journals and research fields were represented in our literature review. It is difficult to expect a researcher who is, for example, interested in evaluating a professional development project with the ATI to know that psychometric evidence of this instrument was also presented in a study seeking to test a model of instructor reflection (Kreber, 2005). But this is precisely why our study, and studies to investigate other popular instruments and inventories, are necessary. At some point, the evidence needs to be gathered in one place so that 20 years after the instrument/inventory has been developed, the research community has a much clearer picture of how valid and reliable the instrument/inventory may be for an individual sample.

We also hope to contribute to growing psychometric standards by presenting limitations of commonly accepted practices and showing the consequences of under-reporting results. Understanding differences between EFA and PCA, how to implement each, and limitations and assumptions of coefficient alpha are all important pieces of presenting evidence of the validity and reliability of data. Also, researchers need to understand the consequences of specific aspects of advanced techniques. For example, it is not surprising that after some researchers included upwards of six error covariance terms in a CFA model, a good fit to the data was achieved. But how does that contribute to the overall construct being measured versus simply inflating fit statistics? In a similar regard, it was a painstaking process trying to figure out the exact factor structures presented in studies due to poor reporting. Most journals today have means for posting supplemental materials, meaning that space constraints are largely not a concern. Every EFA and CFA conducted should report the solutions and appropriate statistics. Even basic information, such as an adequate description of the sample, was missing in several studies. Standard 1.8 of *The Standards* suggests that 'the composition of any sample of test takers from which validity evidence is obtained should be described in as much detail as is practical' (AERA, 2014, p. 25).

As a final note, it is important to remind the reader that we have presented evidence towards the *measurement* of approaches to teaching, not to the actual theoretical basis itself. In other words, no conclusions can be drawn about the theoretical division of approaches to teaching into teacher and student-focused categories. This theory remains an irremovable piece of understanding teaching and learning in university settings and researchers are strongly encouraged to continue investigating it. However, we do hope that the evidence presented here will cause researchers to allocate an appropriate amount of time to identify valid and reliable measures of this construct.

Limitations

In gathering the information for this study, we sometimes had to make educated guesses as to the factor structures proposed by researchers. This leads to the possibility for error, but most factor structures could be determined deductively. As noted earlier, we also recognise that all models based on the ATI16 only have at most 14 items and therefore do not completely represent the original models proposed by their respective authors. Based on the information available in the literature, we would not anticipate our results to change significantly by the addition of these two items in the models analysed.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Science Foundation [NSF CAREER #1552448].

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Baker, L. A., Chakraverty, D., Columbus, L., Feig, A. L., Jenks, W. S., Pilarz, M., … Wesemann, J. L. (2014). Cottrell scholars collaborative new faculty workshop: Professional development for new chemistry faculty and initial assessment of Its efficacy. *Journal of Chemical Education*, *91*(11), 1874–1881. https://doi.org/10.1021/ed500547n

Barbera, J. (2016, August). *Is Cronbach's alpha the correct measure of internal consistency reliability for your data?* Presentation presented at the bienniel conference on chemical education, Greeley, CO.

Chen, J., & Brown, G. T. L. (2016). Tensions between knowledge transmission and student-focused teaching approaches to assessment purposes: Helping students improve through transmission. *Teachers and Teaching*, *22*(3), 350–367.

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

Goh, P. S. C., Wong, K. T., & Hamzah, M. S. G. (2014). The approaches to teaching inventory: A preliminary validation of the Malaysian translation. *Australian Journal of Teacher Education*, *39*(1), 16–26.

Graham, J. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement*, *66*(6), 930–944.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Hativa, N., & Birenbaum, M. (2000). Who prefers what? Disciplinary differences in students' preferred approaches to teaching and learning styles. *Research in Higher Education*, *41*(2), 209–236.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, *30*, 141–158.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55.

Kember, D., & Kwan, K.-P. (2000). Lecturers' approaches to teaching and their relationship to conceptions of good teaching. *Instructional Science*, *28*(5), 469–490.

Åkerlind, G. S. (2004). A new dimension to understanding university teaching. *Teaching in Higher Education*, *9*(3), 363–375. https://doi.org/10.1080/1356251042000216679

Kreber, C. (2005). Reflection on teaching and the scholarship of teaching: Focus on science instructors. *Higher Education*, *50*(2), 323–359.

Lindblom-Ylänne, S., Trigwell, K., Nevgi, A., & Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education*, *31*(3), 285–298.

Martin, E., Prosser, M., Trigwell, K., Ramsden, P., & Benjamin, J. (2002). What university teachers teach and how they teach it. In N. Hativa, & P. Goodyear (Eds.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 103–126). Dordrecht: Springer.

Meyer, J. H. F., & Eley, M. G. (2006). The approaches to teaching inventory: A critique of its development and applicability. *British Journal of Educational Psychology*, *76*(3), 633–649.

Montenegro Maggio, H., & González Ugaldeb, C. (2013). Análisis factorial confirmatorio del cuestionario: 'enfoques de docencia universitaria' (approaches to teaching inventory, ATI-R). *Estudios Pedagógicos (Valdivia)*, *39*(2), 213–230.

Norton, L., Richardson, T. E., Hartley, J., Newstead, S., & Mayes, J. (2005). Teachers' beliefs and intentions concerning teaching in higher education. *Higher Education*, *50*(4), 537–571.

Postareff, L., & Lindblom-Ylänne, S. (2008). Variation in teachers' descriptions of teaching: Broadening the understanding of teaching in higher education. *Learning and Instruction*, *18*(2), 109–120. https://doi.org/10.1016/j.learninstruc.2007.01.008

Preacher, K., & MacCallum, R. (2003). Repairing Tom swift's electric factor analysis machine. *Understanding Statistics*, *2*(1), 13–43.

Prosser, M., & Trigwell, K. (2006). Confirmatory factor analysis of the approaches to teaching inventory. *British Journal of Educational Psychology*, *76*(2), 405–419.

Prosser, M., Trigwell, K., & Taylor, P. (1994). A phenomenographic study of academics' conceptions of science learning and teaching. *Learning and Instruction*, *4*(3), 217–231.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Riggs, I. M., & Enochs, L. G. (1990). Toward the development of an elementary teacher's science teaching efficacy belief instrument. *Science Education*, *74*(6), 625–637.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling | rosseel | journal of statistical software. *Journal of Stastical Software*, *48*(2), 1–36.

Sadler, I. (2009). Development of new teachers in higher education: Interactions with students and other influences upon approach to teaching (dissertation). The University of Edinburgh, Edinburgh. Retrieved from https://www.era.lib.ed.ac.uk/handle/1842/3388

Sadler, I. (2012). The challenges for new academics in adopting student-centred approaches to teaching. *Studies in Higher Education*, *37*(6), 731–745.

Samuelowicz, K., & Bain, J. D. (2001). Revisiting academics' beliefs about teaching and learning. *Higher Education*, *41*(3), 299–325. https://doi.org/10.1023/A:1004130031247

Stains, M., Pilarz, M., & Chakraverty, D. (2015). Short and long-term impacts of the Cottrell scholars collaborative new faculty workshop. *Journal of Chemical Education*, *92*, 1466–1476.
Stes, A., De Maeyer, S. D., & Van Petegem, P. V. (2010). Approaches to teaching in higher education: Validation of a Dutch version of the approaches to teaching inventory. *Learning Environments Research*, *13*(1), 59–73.
Trigwell, K., Prosser, M., & Ginns, P. (2005). Phenomenographic pedagogy and a revised approaches to teaching inventory. *Higher Education Research & Development*, *24*(4), 349–360. https://doi.org/10.1080/07294360500284730
Trigwell, K., & Prosser, M. (1996). Congruence between intention and strategy in university science teachers' approaches to teaching. *Higher Education*, *32*(1), 77–87.
Trigwell, K., & Prosser, M. (2004). Development and use of the approaches to teaching inventory. *Educational Psychology Review*, *16*(4), 409–424.
Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education*, *37*(1), 57–70.
Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology*, *7*, 239. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4880791/
UCLA: Statistical Consulting Group. (n.d.). *Annotated SPSS output: Factor analysis*. Retrieved August 11, 2016, from http://www.ats.ucla.edu/stat/spss/output/factor1.htm
Woolley, S. L., Benjamin, W.-J. J., & Woolley, A. W. (2004). Construct validity of a self-report measure of teacher beliefs related to constructivist and traditional approaches to teaching and learning. *Educational and Psychological Measurement*, *64*(2), 319–331.