

Preliminary results for the development and deployment of Conceptual Learning Assessment Instruments Methodology Survey (CLAIMS)

Julia L. Henning^{1,3}, Kerrie A. Douglas² and Rebecca S. Lindell³

¹*Carthage College, Physics and Astronomy
2001 Alford Park Drive, Kenosha, WI 53140*

²*Purdue University, Engineering Education
701 W Stadium Ave, West Lafayette, IN 47907*

³*Purdue University, Physics and Astronomy
525 Northwestern Ave, West Lafayette, IN 47907*

Abstract. Following the creation of the Force Concept Inventory (FCI), many STEM discipline-based education researchers developed their own version of concept inventories. To incorporate all types of concept inventories created, we introduce the terminology of Conceptual Learning Assessment Instruments (CLAIs). Previous research shows much variation between what is considered a CLAI and the type of evidence used to support the inferences made from the results of CLAIs. As part of our study, we began by creating the Conceptual Learning Assessment Inventory Methodology Survey (CLAIMS). The CLAIMS was sent to developers of over 100 CLAIs identified via a systematic/ structured literature review. This paper discusses the research behind the CLAIMS as well as the preliminary results for the different CLAIs specifically differences in the field-testing of the CLAIs.

PACS: 01.40.Fk, 01.40.gf, 01.50.Kw

I. INTRODUCTION

The creation of the Force Concept Inventory in 1992 and its revision in 1995 [1] revolutionized physics education research. With the ability to quantitatively measure student conceptual understanding, Hake's groundbreaking study shows how interactive engagement methods increased effectiveness in teaching mechanics when compared to traditional teaching methods [2]. Seeing this as a powerful tool for evaluating student learning, discipline-based education researchers across the country began creating similar instruments. Almost twenty-five years after the FCI, there are now over a hundred instruments in the different STEM fields. The necessity of quantitatively measuring student success in a number of contexts has led to a diverse group of instruments across fields.

Previous research [3] shows that developers use different methodologies and have different definitions of a concept inventory. While this variety may not be problematic, it may affect the information that can be validly gained from a particular instrument. Validity must be considered as "nothing less than an evaluative summary of both the evidence for, and the actual – as well as potential – consequences of score interpretation and use" [4]. It is for this reason that we bring attention to the development methodologies and evidence gained to establish validity and reliability in order to compare a wide variety of instruments. For the purposes of this analysis we are adopting a general term to encompass all definitions of

conceptual assessments¹: Conceptual Learning Assessment Instruments (CLAIs).

This study begins to analyze methodologies used to create different STEM CLAIs and the evidence supporting inferences made from their results. In section II, we present the mythologies utilized to identify appropriate CLAIs as well as for the development of the Conceptual Learning Assessment Instrument Methodology Survey (CLAIMS). Sections III presents the results and discussion of the preliminary results of the CLAIMS.

II. METHODOLOGY

A. Identification of CLAIs

We utilized a detailed literature review to identify as many conceptual instruments as possible and then determined the inclusion criteria for an instrument to be considered a CLAI, see Table 1.

TABLE 1. Inclusion criteria CLAIs

Assess content knowledge; Be a coherent instrument ¹ ; Be created for a STEM field; Be published or completed; Be distracter-driven; Be intended for use at the post-secondary level.
¹ We define a coherent instrument to be one created with intent to be used together in the order in which they appear.

¹ The term concept inventory is used by the PER community to refer to an instrument assessing just one concept, while for Engineering Education the term refers to instruments assessing concepts covered in an entire course.

Starting with instruments listed on PhysPort, an AAPT supported website which has previously collected assessment instruments in Physics and Astronomy, researchers identified 37 different CLAI in Physics, Math and Engineering. By searching through the references of these collected publications, more potential publications were identified. Researchers repeated this process until this process yielded no unrecorded publications.

B. Creation of the CLAIMS

The CLAIMS is a 28-item survey, designed to probe the methodologies used to develop the CLAI we collected during our search. A concept inventory expert (Lindell) and a psychometrics expert (Douglas) developed the CLAIMS instrument based on relevant psychometric theory. [4,5] It was sent to CLAI developers previously identified. The CLAIMS asks for developers to provide information on the instrument development methodology, proof of the evidence for the instruments' validity as well as any evidence gathered that was not published at the time of development.

III. RESULTS AND DISCUSSION

A. Preliminary Analysis of CLAIMS

Researchers identified 108 CLAI shown in Fig. 1 in blue. Developers of the identified CLAI were contacted via email and asked to fill out the CLAIMS instrument within 10 days. The developers of only 29 CLAI completed the CLAIMS prior to the deadline as shown in gray in Fig. 1. See Table 2 for the individual CLAI.

1. Differences in instrument design

We see some differences in CLAI design in Table 2. Most CLAI focus on one main concept, but eight of the

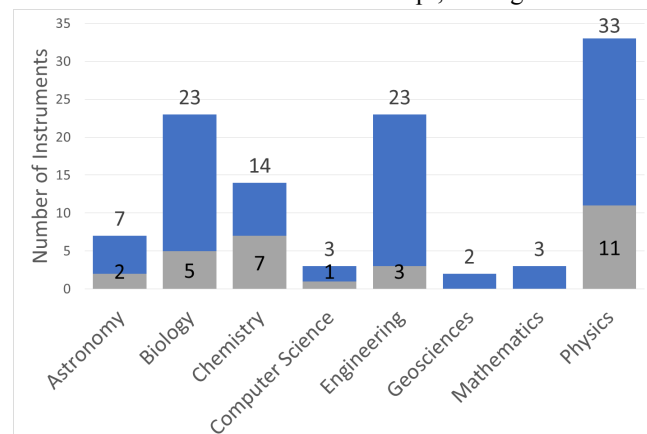


FIG 1: Distribution of CLAI by field. Blue denotes number of CLAI identified. Gray denotes number of CLAIMS instrument submitted.

instruments focus on more than one concept. This is one example of what has necessitated the general classification for these instruments as CLAI and not simply concept inventories.

Table 2. Instruments where developers completed CLAIMS.

Instrument	# of items	# of concepts
Acid-Base Reactions Concept Inventory (ABCI) [6]	28	1
ACID I [7]	9	1
Astronomy Diagnostic Test (ADT) [8]	21	10
Biological Experimental Design Concept Inventory (BEDCI) [9]	14	1
Biology Concept Inventory (BCI) [10]	30	Many
Bonding Representations Inventory (BRI) [11]	23	1
Colorado Classical Mechanics and Math Methods Instrument (CCMI) [12]	11	
Conceptual Inventory of Natural Selection (CINS) [13]	20	1
Colorado Upper-Division Electrostatics assessment (CUE) [14]	17	2
Digital Logic Concept Inventory (DLCI) [15]	24	4
Electric Circuits Concept Inventory (ECCI) [16]	25	Many
Enzyme-Substrate Interactions Concept Inventory (ESICI) [17]	15	1
Flame Test Concept Inventory (FTCI) [18]	19	1
Genetic Drift Inventory (GeDI) [19]	22	1
Host Pathogen Interaction Concept Inventory (HPI-CI) [20]	17	13
Lunar Phases Concept Inventory (LPCI) [21]	20	1
Materials Concept Inventory (MCI) [22]	30	1
Meiosis Concept Inventory (Meiosis CI) [23]	17	4-5
Mechanical Wave Conceptual Survey (MWCS) [24]	22	1
Newtonian Gravity Concept Inventory (NGCI) [25]	26	1
Quantum Mechanics Concept Assessment (QMCA) [26]	31	1
Quantum Mechanics Concept Inventory (QMCI) [27]	9	1
Quantum Mechanics Conceptual Survey (QMCS) [28]	12	
Representational Variant of the FCI (RFCI) [29]	27	1
Relativity Concept Inventory (RCI) [30]	23	8
Redox Concept Inventory (ROXCI) [31]	18	1
Signals and Systems Concept Inventory (SSCI) [32]	25	1
Test of Understanding of Vectors (TUV) [33]	22	1
Vector Knowledge Test (VKT) [34]	9	1

Table 3. Field-testing comparison for CLAIMS responses.

	Appropriate level					Field Testing																						
	High School	Intro Undergrad	Upper Undergrad	Graduate	Other	Age of population					Size				Where			Type of Institution										
						High School	Intro Undergrad	Upper	Graduate	Other	N<100	100 ≤ N ≤ 500	500 ≤ N ≤ 1000	N ≥ 1000	Your Institution	Nearby	Across US	International	High School	Com. College	Small	Small College	Medium College	Large College				
ABCI [7]																												
ACID I [8]																												
ADT [9]																												
BEDCI [10]																												
BCI [11]																												
BRI [12]																												
CCMI [13]																												
CINS [14]																												
CUE [15]																												
DLCI [16]																												
ECCI [17]																												
ESICI [18]																												
FTCI [19]																												
GeDI [20]																												
HPI-CI [21]																												
LPCI [22]																												
MCI [23]																												
Meiosis [24]																												
MWCS [25]																												
NGCI [26]																												
QMCA [28]																												
QMCI [29]																												
QMCS [30]																												
R-FCI [31]																												
RCI [32]																												
ROXCI [33]																												
SSCI [34]																												
TUV [35]																												
VKT [36]																												

2. Differences in field-testing

Because of the limited scope of this paper, we focused on one piece of the CLAI development methodology, the field-testing of the instruments, as this is one component key to validity arguments. Table 3 shows, at the instrument-level, the responses for field-testing of different CLAI. The first section shows the level at which a CLAI should be administered according to the developers, while section two shows the level at which the CLAI was field-tested.

There is discrepancy for some instruments in these two categories. As discussed in previous works [35] it is not

enough to consider classical notions of validity and claim a valid instrument. The evidence collected for their use and interpretation of results supports valid uses of a CLAI. Collecting a sufficient amount of this evidence in field-testing involves both the intended use of the CLAI as well as the size and demographics of the field-testing.

Please note the differences in sections 3, 4, and 5 of Table 3. Considering location of field-testing, 28 CLAI were field tested at the developing institution, but only 18 CLAI were field tested at a location other than the developing institution. For population size most instruments were field-tested on a population of 500 or

fewer (23 CLAI) with almost half of those field-testing on fewer than 100 individuals (11 CLAI). Only 12 CLAI were field tested on more than 500 individuals. We recognize two of CLAI field-testing were for upper-level courses, which often have a smaller population.

With the initial analysis of the CLAIMS, we see a discrepancy with some CLAI between the populations for which the instrument was intended and those on which the instrument was field-tested. Often we see field-testing at the intro-undergraduate level only, likely because of the large easily accessible population in those courses, but then the developers claim that the instrument was developed for use in other age groups. Only 52% of the reported CLAI field- tested on all of their intended age groups that they say the instrument is appropriate to use.

IV. CONCLUSIONS

This study reveals an important issue with the development and use of CLAI, specifically that any use of a CLAI that cannot be supported by appropriate field-testing is not a valid use. Creators of CLAI should be aware they cannot simply establish validity for one population and then claim it is valid for other populations. It is important that evidence for validity and reliability are not bypassed by developers nor overlooked by users so as informed decisions can be made about the reliable and valid uses of the CLAI.

- [1] D. Hestenes, M. Wells, and G. Swackhamer, *Phys. Teach.* **30**, 141 (1992).
- [2] R. Hake, *Am. J. Phys.* **1** (1998).
- [3] R. S. Lindell, E. Peak, and T. M. Foster, in *AIP Conf. Proc.* (AIP, 2007), pp. 14–17.
- [4] S. Messick, *Am. Psychol.* **50**, 741 (1995).
- [5] T. Raykov, *Introduction to Psychometric Theory* (Routledge, New York, 2011).
- [6] J. D. Jensen, Miami University, Ph.D. Dissertation, May 2013
- [7] L. M. McClary and S. L. Bretz, *Int. J. Sci. Educ.* **34**, 2317 (2012).
- [8] B. Hufnagel, *Astron. Educ. Rev.* **1**, 47 (2001).
- [9] T. Deane, K. Nomme, E. Jeffery, C. Pollock, and G. Birol, *Cell Biol. Educ.* **13**, 540 (2014).
- [10] K. Garvin-Doxas and M. Klymkowsky, Unpubl. Manusc. That Can Be Found [Http//bioliteracy.net/CABS](http://bioliteracy.net/CABS) **202007**, 1829 (2007).
- [11] C. J. Luxford and S. L. Bretz, *J. Chem. Educ.* **91**, 312 (2014).
- [12] M. D. Caballero and S. J. Pollock, in *2013 Phys. Educ. Res. Conf. Proc.* (2014), pp. 81–84.
- [13] D. L. Anderson, K. M. Fisher, and G. J. Norman, *J. Res. Sci. Teach.* **39**, 952 (2002).
- [14] S. V. Chasteen, R. E. Pepper, M. D. Caballero, S. J. Pollock, and K. K. Perkins, *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **8**, 1 (2012).
- [15] G. L. Herman, M. Loui, and C. Zilles, *Proc. 41st ACM Tech. ...* 102 (2010).
- [16] T. Ogunfunmi and M. Rahman, in *ISCAS 2010 - 2010 IEEE Int. Symp. Circuits Syst. Nano-Bio Circuit Fabr. Syst.* (2010), pp. 2804–2807.
- [17] S. L. Bretz and K. J. Linenberger, *Biochem. Mol. Biol. Educ.* **40**, 229 (2012).
- [18] S. L. Bretz, K. J. Linenberger, *Biochemistry and Molecular Biology Education*, **40**(4), 229-233 (2012).
- [19] R. M. Price, T. C. Andrews, T. L. McElhinny, L. S. Mead, J. K. Abraham, A. Thanukos, and K. E. Perez, *CBE - Life Sci. Educ.* **13**, 65 (2014).
- [20] G. Marbach-Ad, V. Briken, N. M. El-Sayed, K. Frauwirth, B. Fredericksen, S. Hutcheson, L.-Y. Gao, S. W. Joseph, V. Lee, K. S. McIver, D. Mosser, B. Booth Quimby, P. Shields, W. Song, D. C. Stein, R. T. Yuan, and A. C. Smith, *J. Microbiol. Biol. Educ.* **10**, 43 (2009).
- [21] R. Lindell and J. Olsen, *Proc. 2002 Phys. Educ. Res. Conf.* (2002).
- [22] S. Krause, J. C. Decker, and R. Griffin, *33rd Annu. Front. Educ. 2003. FIE 2003.* **1**, 7 (2003).
- [23] P. Kalas, A. O'Neill, C. Pollock, and G. Birol, *CBE Life Sci. Educ.* **12**, 655 (2013).
- [24] A. Tongchai, M. D. Sharma, I. D. Johnston, K. Arayathanikul, and C. Soankwan, *Int. J. Sci. Educ.* **31**, 2437 (2009).
- [25] K. E. Williamson, S. Willoughby, E. F. Prather, *Astronomy Education Review* v. 12, n. 1 (2013)
- [26] H. R. Sadaghiani and S. J. Pollock, *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **11**, 1 (2015).
- [27] J. Falk, johan@vaxjonexus.com (2015).
- [28] S. B. McKagan, K. K. Perkins, and C. E. Wieman, *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **6**, 1 (2010).
- [29] P. Nieminen, A. Savinainen, and J. Viiri, *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **6**, 1 (2010).
- [30] J. S. Aslanides and C. M. Savage, *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **9**, 1 (2013).
- [31] A. R. Brandriet and S. L. Bretz, *J. Chem. Educ.* **91**, 1132 (2014).
- [32] K. E. Wage and J. R. Buck, *31st Annu. Front. Educ. Conf. Impact Eng. Sci. Educ. Conf. Proc. (Cat. No.01CH37193)* **2**, (2001).
- [33] P. Barniol and G. Zavala, *Physical Review Special Topics- Physics Education Research*, **10**, 0101215 (2013).
- [34] R. D. Knight, *Phys. Teach.* **33**, 74 (1995).
- [35] K. A. Douglas and Ş. Purzer, *J. Eng. Educ.* **104**, 108 (2015).