Virtual Tools for Developing SCLAs: An Example of Concentration Analysis

Andrew J. Mason,[1] Liam G. E. McDermott,[2] Mayuri Gilhooly,[3] Dawn Meredith,[4] James Vesenka,[3] D. J. Wagner,[5] Daniel Young,[6] Rebecca Lindell[7]

[1]Department of Physics, Astronomy, and Engineering, University of Central Arkansas, Conway, AR 72035
[2]Department of Physics, University of Connecticut – Avery Point, Groton, CT 06340
[3]School of Molecular and Physical Sciences, University of New England, Biddeford, ME 04005
[4]Department of Physics and Astronomy, University of New Hampshire, Durham, NH 03824
[FF5]Department of Physics, Grove City College, Grove City, PA 16127
[6]Department of Physics, Westminster University, Salt Lake City, UT 84105
[7]Tiliadal Solutions, Lafayette, IN 47901

Abstract

Standardized Conceptual Learning Assessments (SCLAs) measure individuals' conceptual understanding using distractor-driven multiple-choice items. Concentration analysis (CA), a quantitative technique that evaluates the concentrations of each item's response distribution allowing researchers to determine how many models are triggered by each item [1], and can be utilized to help establish the validity of distractor driven items [2]. As part of the development of the Fluids Conceptual Evaluation (FCE), a distractor-driven two-tier multiple-choice evaluation instrument, the authors evaluated the concentration factors of a set of responses from the FCE's preliminary pilot test results. Results of this analysis show how many different mental models are triggered by each item's first tier. This paper demonstrates how the use of CA verifies recommended models, and suggests a similar model structure for concentration deviation. Recommendations for future use of concentration analysis will be made accordingly.

Background

- Distractor-driven assessments: context may depend on respondent interpretation
- ***Concentration analysis*** (CA) [1, 2]: goes beyond score (correct vs. incorrect) to examine patterns of student responses
- E.g. incorrect answer patterns suggest p-prims [3]

- ***Fluids Conceptual Examination (FCE), beta version***: 59 items grouped in 10 fluid dynamics constructs
- Two-tier format for items: 1st tier = standard MC, 2nd tier asks for reasoning on 1st tier
  - (*Focus on 1st tier only here*)
- Long-term goals: Assist IPLS instructors with introducing fluids topics; determine CA's role in developing SCLAs

Research questions:

1. *Does CA detect similar patterns for FCE items as with prior literature? [1]*
2. *How do S-Γ models look with respect to S-C models?*
3. *What items can be featured as a preliminary example?*

Methodology

- Data collection (de-identified): 259 respondents from 3 institutions, each respondent assigned 8 two-tier items randomly (1 preliminary round)

- Concentration factor C [1,4]: Distribution of all choices (correct and incorrect), 0 = even distribution, 1 = everybody picked the same answer

$$C = \frac{\sqrt{m}}{\sqrt{m}-1}\left(\frac{\sqrt{\sum_{i=1}^{m} n_i^2}}{N} - \frac{1}{\sqrt{m}}\right)$$

- Concentration deviation Γ [1]: Distribution of incorrect choices only, 0 = randomized distribution, 1 = all incorrect responses were the same choice

$$\Gamma = \frac{\sqrt{m}}{\sqrt{m}-1}\left(\frac{\sqrt{(\sum_{i=1}^{m} n_i^2) - n_{correct}^2}}{N - n_{correct}} - \frac{1}{\sqrt{m}}\right)$$

- N = # of total responses for item; m = # of item choices; n_i = # of respondents picking i^{th} choice
- S-C and S-Γ curve classifications [2, 5]:
- Score (S): Low (< 0.4), Medium (0.4 – 0.7), High (> 0.7)
- C and Γ: Low (< 0.2), Medium (0.2 – 0.5), High (> 0.5)

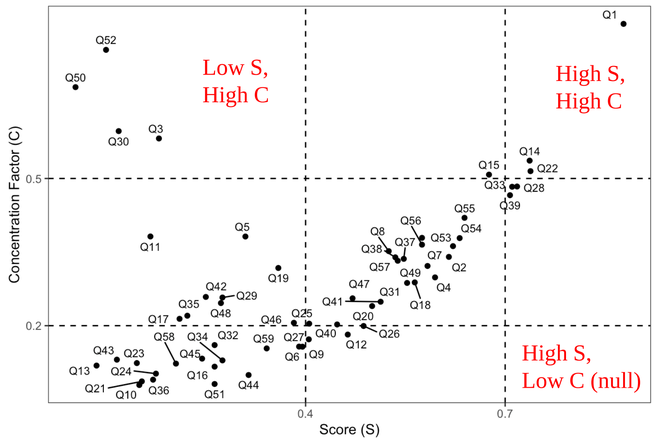

FIG. 1: Score-concentration factor (S-C) plot for items in Table 1.

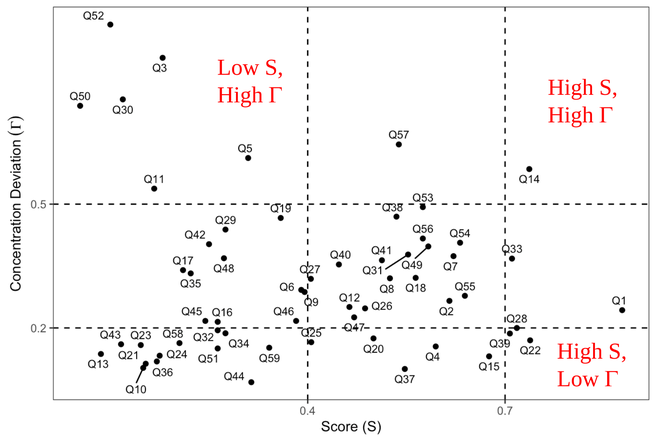

FIG. 2: Score-concentration deviation (S-Γ) plot for Table 1 items.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	41	39	39	37	42	41	37	40	43	40	42	41	35	38	37
m	8	4	4	4	7	4	4	7	4	8	4	4	11	4	7
S	**0.88**	0.62	*0.18*	0.59	*0.31*	*0.39*	0.62	0.53	*0.40*	*0.15*	*0.17*	0.46	*0.09*	**0.74**	0.68
C	**0.82**	0.34	**0.58**	0.30	0.38	*0.16*	0.36	0.35	*0.16*	*0.08*	0.38	*0.18*	*0.12*	**0.54**	**0.51**
G	0.24	0.27	**0.85**	*0.16*	**0.61**	0.29	0.37	0.32	0.29	*0.10*	**0.54**	0.25	*0.14*	**0.59**	*0.13*
S-C	H,H	M,M	L,H	M,M	L,M	L,L	M,M	M,M	L,L	L,L	L,M	M,L	L,L	H,H	M,H
S-Γ	H,M	M,M	L,H	M,L	L,H	L,M	M,M	M,M	L,M	L,L	L,H	M,M	L,L	H,H	M,L
Item	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
N	38	38	39	39	40	39	42	41	40	37	39	42	39	40	42
m	4	4	4	8	5	8	4	5	4	7	4	4	4	5	4
S	*0.26*	*0.21*	0.56	*0.36*	0.50	*0.15*	**0.74**	*0.15*	*0.18*	0.41	0.49	0.40	**0.72**	*0.28*	0.12
C	*0.12*	0.21	0.29	0.32	0.24	*0.09*	**0.51**	*0.12*	*0.10*	0.20	*0.20*	*0.17*	0.48	0.26	**0.60**
G	0.21	0.34	0.32	0.47	*0.17*	*0.11*	*0.17*	*0.16*	*0.13*	*0.17*	0.25	0.32	*0.20*	0.44	**0.75**
S-C	L,L	L,M	M,M	L,M	M,M	L,L	H,H	L,L	L,L	M,M	M,L	M,L	H,M	L,M	L,H
S-Γ	L,M	L,M	M,M	L,M	M,L	L,L	H,L	L,L	L,L	M,L	M,M	M,M	H,L	L,M	L,H
Item	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**
N	38	38	38	40	36	41	42	43	41	38	39	40	43	35	45
m	4	9	4	6	6	8	7	5	4	4	4	6	9	7	5
S	0.55	*0.26*	**0.71**	*0.28*	*0.22*	*0.17*	0.55	0.53	**0.71**	0.45	0.51	***0.25***	*0.12*	*0.31*	*0.24*
C	0.29	*0.16*	0.48	*0.13*	0.22	*0.09*	0.34	0.34	0.47	0.20	0.25	0.26	*0.13*	*0.10*	*0.13*
G	0.38	*0.19*	0.37	*0.19*	0.33	*0.12*	*0.10*	0.47	*0.19*	0.35	0.36	0.40	*0.16*	*0.07*	0.22
S-C	M,M	L,L	H,M	L,L	L,M	L,L	M,M	M,M	H,M	M,M	M,M	L,M	L,L	L,L	L,L
S-Γ	M,M	L,L	H,M	L,L	L,M	L,L	M,L	M,M	H,L	M,M	M,M	L,M	L,L	L,L	L,M
Item	**46**	**47**	**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	
N	34	34	44	36	37	38	40	40	38	36	40	39	39	41	
m	7	6	7	4	4	4	4	5	4	5	5	4	8	7	
S	*0.38*	0.47	*0.27*	0.58	*0.05*	*0.26*	*0.10*	0.58	0.63	0.64	0.58	0.54	*0.21*	*0.34*	
C	0.21	0.26	0.25	0.32	**0.69**	*0.08*	**0.76**	0.38	0.38	0.42	0.37	0.33	*0.12*	*0.15*	
G	0.22	0.23	0.37	0.40	**0.74**	*0.15*	**0.94**	0.49	0.41	0.28	0.42	**0.64**	*0.16*	*0.15*	
S-C	L,M	M,M	L,M	M,M	L,H	L,L	L,H	M,M	M,M	M,M	M,M	M,M	L,L	L,L	
S-Γ	L,M	M,M	L,M	M,M	L,H	L,L	L,H	M,M	M,M	M,M	M,M	M,H	L,L	L,L	

TABLE I: Number of responses (N), number of item choices (m), score (S), concentration factor (C), concentration deviation (Γ), S-C type (S-C), and S-Γ type (S-Γ) for each of the 59 preliminary items for the FCE within the student sample. Based on recommended scales [1,2], "high" values (H) are in **bold**, "medium" values (M) are in normal font, and "low" values (L) are in *italics*. The patterns for L, M, and H values are presented in the S-C row in the form of "S level, C level," and similarly for the S-Γ row in the form of "S level, Γ level."

Results

- ***Key difference from Refs. 1, 2:*** number of choices (m) differs between items
- **Table 1: How do 59 items fit prescribed S-C model types?**
  - 28 items in "two-model" type
    - 19 items: correct answer + 1 distractor
    - 9 items: 2 distractors
  - 17 items in "non-model" type (3+ choices)
  - 7 items in "one-model" type
    - 4 items: correct answer dominant
    - 3 items: specific distractor dominant
  - 7 items in between model types
- **Fig. 1:** C_{max}, C_{min} bounds [1] not as well-defined due to unequal m values
- **Fig. 2: S-C model types seem to work for S-Γ**
  - Differences: Γ decoupled from correct responses

Initial Conclusions

- ***Most items fit within S-C model types*** [1]
  - Other items in "between-model" state
- S-Γ graph: ***most items also fit S-C model types***
  - Lower S = higher overlap with Fig. 1
  - ***Moderate and high scores: S-Γ graph more useful as separate measure***
- p-prims [3] clearly identified with low-S, high-C/Γ items (strong distractor choice)

Future Work

- ***Does this analysis work for Tier 2 items?***
  - Tier 2 responses depend on Tier 1 responses
- Closer look within each construct: find minimum orthonormal set of items
- Eventually finalize FCE to approximately 30 items
- Eqns. 1-2: Tweak for varying m to fix C_{min}, C_{max}

References

1. Bao, L., & Redish, E. F. (2001). Concentration analysis: A quantitative assessment of student states. American Journal of Physics, 69(S1), S45-S53.
2. Lindell, R. S., & Olsen, J. P. (2002). Developing the Lunar Phases Concept Inventory. In 2002 PERC Proceedings [Boise, ID, August 7-8, 2002], edited by S. Franklin, K. Cummings, and J. Marx, doi:10.1119/perc.2002.pr.011. <https://www.compadre.org/Repository/document/ServeFile.cfm?ID=4323&DocID=1162>.
3. diSessa, A. A. (1993). Toward an Epistemology of Physics. Cognition and Instruction, 10(2–3), 105–225.
4. Wang, J., & Bao, L. (2010). Analyzing force concept inventory with item response theory. American Journal of Physics, 78(10), 1064-1070.
5. Dega, B. G., & Govender, N. (2016). Assessment of students' scientific and alternative conceptions of energy and momentum using concentration analysis. African Journal of Research in Mathematics, Science and Technology Education, 20(3), 201-213.

Acknowledgements

- NSF Award # 2021059 (AAPT)
- NSF Award # 2021261 (UNE)
- NSF Award # 2021273 (UNH)
- Participating universities and instructors for data collection
- UCA Department of Physics, Astronomy, and Engineering
- Faculty Development Grant, UCA Center for Excellence in Teaching And Learning