

The Score



DIVISION 5 PERIODICAL VOLUME 47 ISSUE 3

EDITOR: GILL SITARENIOS

ASSISTANT EDITOR: JOANNA SOLOMON



In this issue . . .

ZAFAR M. IQBAL, RODICA IVAN, KAITLIN BYNKOSKI, KRISTEN ARCHBELL, CYNTHIA RICHARD, & KRISTINA H. PETERSEN	2
Formative SJT: Developing Student Personal and Professional Competencies	
MEAGAN DOCHERTY, ALEXIS CHAVEZ, & TIANYING SUN	11
Division 5 Mentoring Program	
ELIZABETH A. SANDERS	13
Division 5 APA Sessions and Events	
ALSO IN THIS ISSUE	14
What's New in Books (p. 14); New Editor (p. 16); Links to previous issues (p. 17); Division 5 Membership Information (p. 18); Advertising in The Score (p. 19)	

FORMATIVE SJT: DEVELOPING STUDENT PERSONAL AND PROFESSIONAL COMPETENCIES

*Zafar Iqbal, Ph.D., Rodica Ivan, Ph.D., Kaitlin Bynkoski, PharmD,
Kristen Archbell, Ph.D., Cynthia Richard, Ph.D., & Kristina Petersen, Ph.D.*



Zafar Iqbal



Rodica Ivan



Kaitlin Bynkoski



Kristen Archbell



Cynthia Richard



Kristina Petersen

Personal and professional skills (i.e., empathy, collaboration, communication, ethics etc.) are integral to success in both academic and professional settings. These competencies not only enhance healthcare professionals' academic performance but also contribute to long-term career development (McGill, Ali, & Barton, 2020; Saunders & Bajjaly, 2022). Despite their significance, health professions training programs often face challenges in effectively assessing and nurturing these skills due to limited resources, lack of standardized evaluation tools and personalized feedback (Alt, Naamati-Schneider, & Weishut, 2023; Saunders & Bajjaly, 2022). To address this gap, formative situational judgment tests (SJTs) offer a promising solution. SJTs provide structured, context-rich scenarios that can evaluate and develop essential soft skills in a low-stakes, feedback-driven manner (Patterson et al., 2019; Sahota et al., 2023; Cullen et al., 2022), and can help identify students in need of early intervention and professional growth support (Cullen et al., 2017, 2020).

In this paper, we outline the design and utility of a formative SJT assessment, ACE (Acuity Competency Enhancer), referencing its application in two health sciences programs. Within the context of this tool, social and personal intelligence is understood as the ability to critically reflect on and respond to interpersonal and professional challenges.

Background

SJTs in health professions education can be implemented across the learning spectrum or continuum—from admissions and undergraduate training to postgraduate selection and ongoing professional development. SJTs have been delivered in various formats (e.g., free-text, multiple choice type, video recordings, etc.) and used for both formative and summative purposes (Ballejos et al., 2024; Lievens & Motowidlo, 2016; Patterson & Driver, 2018; Saxena et al., 2024).

A recent systematic literature review uncovered eight studies that utilized bespoke, in-house developed SJTs predominately as formative assessments (Foucault et al. 2015; Kiessling et al. 2016; Frohlich et al. 2017; Goss et al. 2017; Antes et al. 2020; Graupe et al. 2020; Ludwig et al. 2021; Reiser et al. 2021). Findings from these studies collectively suggest potential links between SJT performance and professional behaviors. For instance, traits like agreeableness, conscientiousness, extraversion, and openness were associated with stronger performance, while neuroticism predicted lower scores. Some studies (Antes et al., 2020; Foucault et al., 2015; Goss et al., 2017) used SJTs as learning tools, though without measuring knowledge or attitude gains. To

support students' development of the measured skills, Goss et al. (2017) added feedback for low performers, an approach echoed in other studies. These studies collectively align with broader thinking about professional identity formation and suggest formative SJTs may enhance learning. Noteworthy, all formative SJT studies used close-ended, text-based scenarios with ranking or multiple-choice formats, though video-based delivery (Frohlich et al., 2017; Graupe et al., 2020; Reiser et al., 2021) also showed promise for fostering desirable traits.

Assessment Design

Acuity Competency Enhancer (ACE) was developed as a formative, open-response SJT combining both qualitative and quantitative measures, and provides feedback to all test takers to help them improve their personal and professional skills. The formative SJT version used in this study included 13 scenarios. Open-ended responses were rated on a scale of 1 to 9 by experienced raters trained on ACE-specific scoring criteria.

ACE comprises four domains; each domain reflects a distinct yet interrelated skill area as depicted in Figure 1. These domains were created as aggregated composites as a series of sub-constructs that were reflected in the question content of the SJT. Specifically, the "READ" domain comprises empathy, equity, and ethics. The "YOU" dimension comprises self-awareness, resilience, and motivation. The "INTER" dimension comprises communication and collaboration. And, finally, the "ENACT" dimension comprises problem solving and professionalism. Together, these domains enable a holistic assessment of personal and professional skills, potentially informing both individual student development, broader educational strategies and curricular reforms.

Figure 1. ACE Four Domain Model.

Reading the Situation (READ): Accurately interpreting context and nuance Concepts: ethics, active listening, body language, compassion, cultural sensitivity, bias, client safety	How it Affects You (YOU): Recognizing personal reactions, experiences, and biases Concepts: Self-awareness, motivation, accountability, responsibility, self-evaluation, life-long learning, honesty
Effective Interacting (INTER): Engaging others constructively and respectfully, navigating social dynamics Concepts: Communication, collaboration, teamwork, interpersonal skills, empathy	Resolution Enactment (ENACT): Implementing thoughtful, adaptable solutions to dilemmas, demonstrating commitment and energy to see them through Concepts: problem solving, critical thinking, persistence, recognition, contingency planning, assumption analysis

Initial Psychometric Indicators: Reliability

Initial technical reports (Sitarenios, 2024a; 2024b) provided early evidence supporting its reliability and construct validity. For the open-ended items, domain-level reliability estimates ranged from .79 to .92. In contrast, reliability for the aggregated multiple-choice score was lower, at .63. A four-factor model corresponding to the ACE domains demonstrated acceptable model fit (CFI = .96, TLI = .93, RMSEA = .12).

This report draws on aggregate data (N = 338) from two programs who participated in the pilot phase of the development of this formative SJT. A total of 219 students included came from students from the Class of 2027

from New York Medical College's School of Medicine MD program, and an additional 119 second year students from University of Waterloo, School of Pharmacy program. A demographic breakdown of the entire sample is shown in Table 1.

Table 1. Demographic breakdown of students included in this study.

Gender	Percentage
Male	36.1
Female	63.2
Other	0.6
Ethnicity	Percentage
Black, African, Caribbean, or African American	5.6
East Asian (Including Chinese, Japanese, Korean, Mongolian, Tibetan, and Taiwanese)	18.5
Hispanic, Latinx, or Spanish origin	5.6
Middle Eastern or Northern African	9.3
South Asian (Including Bangladeshi, Bhutanese, Indian, Nepali, Pakistani, and Sri Lankan)	16.9
White or European	37.7
Other	6.0

Using these data, both internal consistency reliability (Table 2) and inter-rater reliability (Table 3) was examined. Internal consistency is reported using both Cronbach's *alpha* and McDonald's *omega*. While *alpha* remains the most frequently cited metric, it assumes item homogeneity and may underestimate reliability under certain conditions (Deng & Chan, 2017). *Omega* is considered more robust, especially for multidimensional constructs, as it imposes fewer assumptions and often provides a more accurate estimate (McDonald, 1999). In this analysis, all four domain scores showed strong internal consistency, with *alpha* values ranging from .85 to .89 and *omega* values from .88 to .92 (see Table 2).

Table 2. ACE reliability by domain score.

	# items	Alpha	Omega	Median Item <i>r</i>
READ - OE	7	.86	.90	.49
YOU - OE	7	.85	.88	.45
INTER - OE	10	.89	.92	.45
ENACT - OE	8	.88	.91	.44

Inter-rater reliability data were available for a subset of the data. All scenarios were rated by a second distinct set of raters for a random group of 50 of the students. These paired ratings were then used to compute IRR statistics. When aggregated by scale, all four dimensions demonstrated strong inter-rater reliability, supporting consistent scoring practices for this formative SJT (see Table 3).

Table 3. Inter-rater reliability statistics for R1 rater group for overall domain scores.

Metric	R1 ¹
READ - OE	.94
YOU - OE	.95
INTER - OE	.92
ENACT - OE	.92

¹All correlations significant, $p < .001$

Item-level rater consistency was with 35 independent questions being rated by two independent raters, and inter-rater correlations were calculated for each item. Strong inter-rater consistency was observed at the item level, further supporting the reliability of individual question scores (see Table 4).

Table 4. Question level Inter-Rater Reliability.

	Average	Minimum	Maximum	SD
Pearson r	.79	.63	.91	.07

In-program utilization

Beyond basic psychometric evaluation, a key goal of this investigation was to explore the tool's utility in supporting programs in measuring and developing students' personal and professional skills. The following sections describe its implementation in two pilot programs.

Program A (New York Medical College)

Students in this program completed the ACE assessment at the end of their first semester, integrated into their coursework as part of their *Physicianship and Professionalism* module. They received formative feedback on their open-ended responses for two scenarios. Figure 2 illustrates an example of the feedback provided.

Figure 2. Sample Scenario Feedback.

Scenario Content:

This scenario presents a dilemma where the applicant must navigate the ethical problem of whether to accept a well-paying position at a big firm where covering up unethical practices may be necessary. It is unclear if the concerns about unethical practices are legitimate or if they are based on rumors. The company is also requiring a non-disclosure as part of the contract to be signed to accept the job.

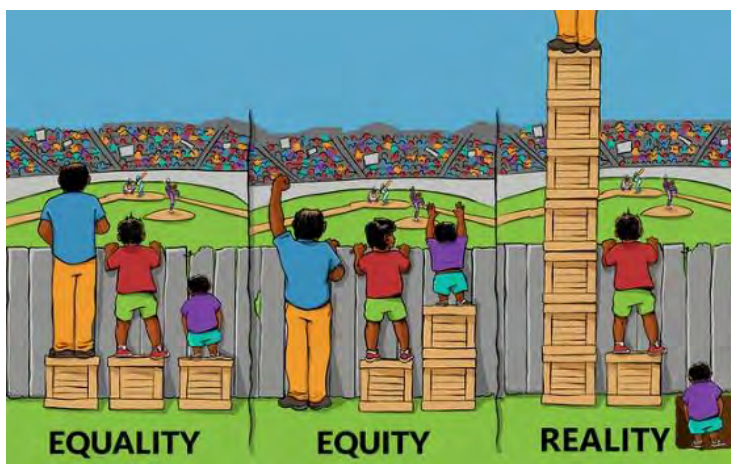
Sample Feedback:

Your responses show a good start in considering multiple perspectives, particularly in the first question where you acknowledge Mel's lack of prior mistakes and suggest that education might be a more appropriate response than termination. For a more comprehensive perspective, it would be beneficial to also consider the perspectives of the patron affected by the allergy, the restaurant management, and the potential legal implications for the restaurant. In terms of justification and rationale, your arguments are coherent and clear, but they could be strengthened by providing more depth. For instance, in your second response, you could elaborate on specific collaborative strategies that Mel could use with her coworkers, kitchen staff, and management to address the situation and prevent future incidents. This would show a deeper understanding of the collaborative process and its benefits. In your third response, you've identified potential emotions Mel might be feeling, which is great, but consider also how these emotions could affect her interactions with others and her decision-making process. Overall, you've made a good effort, and with a bit more detail and consideration of all parties involved, your responses could be even more insightful.

Students also received feedback on their performance on the multiple-choice section of the assessment. One item focused on distinguishing between equity and equality. Figure 3 presents this portion of the feedback.

Figure 3. Equity Feedback.

Equality and equity are two distinct concepts, although they're often used interchangeably. Where equality gives everyone access to the same opportunities, equity in the workplace means that access to opportunities is proportional to each person's needs. This ensures that the opportunity is equally advantageous to everybody, because it's offered in a way that takes diversity into account. In other words, equity will eliminate barriers and level the playing field by giving each person the right kind of support. Angus Maguire's popular cartoon (see below) illustrates the difference with a visual metaphor. It shows three people of varying heights standing on boxes to watch sports from behind a fence. In the 'equality' part of the image, each person receives the same number of boxes to stand on, meaning only the tallest can see over the fence. In the 'equity' part of the image, each person has the number of boxes they need to see over the fence, creating an equal advantage for everyone.



Cartoon Attribution: [Interaction Institute for Social Change](#) | **Artist:** [Angus Maguire](#)

Program B (University of Waterloo, School of Pharmacy)

As part of its annual Midpoint Assessment, the School of Pharmacy traditionally administers an OSCE and a multiple-choice exam to identify students who may need remediation. In a recent enhancement, the ACE was added as a third component to better support holistic student development, addressing not only academic performance but also professionalism-related competencies. A remediation framework was implemented based on ACE domain scores, classifying students into three categories:

- No Remediation – Students showing moderate to high competency across all domains (82%)
- Resource Review Recommended – Students with one domain in the “developing” range (10%)
- Advisory Support Recommended – Students with two or more developing domains (7%)
- Remediation with Academic Advisor Strongly Recommended (2%)

Importantly, there was minimal overlap between students flagged by the SJT and those identified through the OSCE/MCQ assessments—only one student required remediation across all three components. This suggests that the ACE SJT effectively identifies a unique subset of students needing support in professionalism competencies.

Discussion

Future development efforts will focus on refining the assessment to improve its psychometric robustness, incorporating feedback from pilot deployments, and further aligning the tool with student support and remediation strategies.

Alt, D., Naamati-Schneider, L., & Weishut, D. J. N. (2023). Competency-based learning and formative assessment feedback as precursors of college students' soft skills acquisition. *Studies in Higher Education*, 48(12), 1901–1917. <https://doi.org/10.1080/03075079.2023.2217203>

Antes, A. L., Dineen, K. K., Bakanas, E., Zahrli, T., Keune, J. D., Schuelke, M. J., & DuBois, J. M. (2020). Professional decision-making in medicine: Development of a new measure and preliminary evidence of validity. *PLoS ONE*, 15(2), e0228450. <https://doi.org/10.1371/journal.pone.0228450>
pubmed.ncbi.nlm.nih.gov+1nottingham-repository.worktribe.com+1

Ballejos, M. P., Cestone, C. Copeland, H. L., Dunleavy, D. M., Geiger, T., & Patel, D. M. S. (2024). Predicting Medical Student Performance with a Situational Judgment Test in Admissions. *Academic Medicine* 99(2),175-182. <https://doi.org/10.1097/ACM.00000000000005516>

Cullen, M. J., Zhang, C., Marcus-Blank, B., Braman, J. P., Tiryaki, E., Konia, M., Hunt, M. A., Lee, M. S., Van Heest, A., Englander, R., & Sackett, P. R. (2020). Improving our ability to predict resident applicant performance: Validity evidence for a situational judgment test. *Teaching and Learning in Medicine*, 32(5), 508–521. <https://doi.org/10.1080/10401334.2020.1760104>

Cullen, M. J., Zhang, C., Sackett, P. R., Thakker, K., & Young, J. Q. (2022). Can a situational judgment test identify trainees at risk of professionalism issues? A multi-institutional, prospective cohort study. *Academic Medicine*, 97(10), 1494–1503. <https://doi.org/10.1097/ACM.00000000000004756>

Cullen, M. J., Konia, M. R., Borman-Shoap, E. C., Braman, J. P., Tiryaki, E., Marcus-Blank, B., & Andrews, J. S. (2017). Not all unprofessional behaviors are equal: The creation of a checklist of bad behaviors. *Medical Teacher*, 39(1), 85–91. <https://doi.org/10.1080/0142159X.2016.1231917>

Deng, L., & Chan, W. (2017). Testing the Difference Between Reliability Coefficients Alpha and Omega. *Educational and psychological measurement*, 77(2), 185–203. <https://doi.org/10.1177/0013164416658325>

- Foucault, A., Dubé, S., Fernandez, N., Gagnon, R., & Charlin, B. (2015). Learning medical professionalism with the online Concordance-of-Judgment learning tool (CJLT): A pilot study. *Medical Teacher*, 37(10), 955–960. <https://doi.org/10.3109/0142159X.2014.970986> md.umontreal.ca+1pubmed.ncbi.nlm.nih.gov+1
- Fröhlich, M., Kahmann, J., & Kadmon, M. (2017). Development and psychometric examination of a German video-based situational judgment test for social competencies in medical school applicants. *International Journal of Selection and Assessment*, 25(1), 94-110.
- Goss, B. D., Ryan, A. T., Waring, J., Judd, T., Chiavaroli, N. G., O'Brien, R. C., Trumble, S. C., & McColl, G. J. (2017). Beyond selection: The use of situational judgment tests in the teaching and assessment of professionalism. *Academic Medicine*, 92(6), 780–784. <https://doi.org/10.1097/ACM.0000000000001591>
- Graupe, T., Fischer, M. R., Strijbos, J. W., & Kiessling, C. (2020). Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Patient education and counseling*, 103(9), 1839–1845. <https://doi.org/10.1016/j.pec.2020.04.001>
- Kiessling, C., Bauer, J., Gartmeier, M., Iblher, P., Karsten, G., Kiesewetter, J., Moeller, G. E., Wiesbeck, A., Zupanic, M., & Fischer, M. R. (2016). Development and validation of a computer-based situational judgement test to assess medical students' communication skills in the field of shared decision making. *Patient education and counseling*, 99(11), 1858–1864. <https://doi.org/10.1016/j.pec.2016.06.006>
- Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology*. 9(1), 3-22. <http://doi.org/10.1017/iop.2015.71>
- Ludwig, S., Behling, L., Schmidt, U., & Fischbeck, S. (2021). Development and testing of a summative video-based e-examination in relation to an OSCE for measuring communication-related factual and procedural knowledge of medical students. *GMS journal for medical education*, 38(3), Doc70. <https://doi.org/10.3205/zma001466>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- McGill, C. M., Ali, M., & Barton, D. (2020). Skills and competencies for effective academic advising and personal tutoring. *Frontiers in Education*, 5, 135. <https://doi.org/10.3389/feduc.2020.00135>
- Patterson F., & Driver R. (2018). Situational judgement tests (SJTs). In F. Patterson & L. Zibarras (Eds.), *Selection and recruitment in the healthcare professions: Research, theory and practice* (pp.79-112). Palgrave Macmillan.
- Patterson, F., Galbraith, K., Flaxman, C., & Kirkpatrick, C. M. (2019). Evaluation of a situational judgement test to develop non-academic skills in pharmacy students. *American journal of pharmaceutical education*, 83(10), 7074.
- Reiser, S., Schacht, L., Thomm, E., Bauer, J., Figalist, C., Janssen, L., Schick, K., Berberat, P. O., Gartmeier, M., & Dörfler, E. (2021). A video-based situational judgement test of medical students' communication competence in patient encounters: Development and first evaluation. *Patient Education and Counseling*, 105(6), 1283–1289. <https://doi.org/10.1016/j.pec.2021.10.012>

- Sahota, G. S., Fisher, V., Patel, B., JuJ, K., & Taggar, J. S. (2023). The educational value of situational judgement tests (SJTs) when used during undergraduate medical training: A systematic review and narrative synthesis. *Medical Teacher*, 45(9), 997–1004. <https://doi.org/10.1080/0142159X.2023.2168183>
- Saunders, L., & Bajjaly, S. (2022). Direct instruction and assessment of personal and professional skills across disciplines: Faculty perspectives. *International Journal of Teaching and Learning in Higher Education*, 33(3), 374–384.
- Saxena, A., Desanghere, L., Dore, K., & Reither, H. (2024) Incorporating a situational judgment test in residency selections: Clinical, educational and organizational outcomes. *BMC Medical Education* 24(339). <https://doi.org/10.1186/s12909-024-05310-8>
- Sitarenios, G. (2024a). *Phase 2 ACE Pilot Research Analysis - 1*. Technical Report. Acuity Insights.
- Sitarenios, G. (2024b). *Phase 2 ACE Pilot Research Analysis - 2*. Technical Report. Acuity Insights.



Division 5's early career committee and graduate student committee are pleased to announce that the Division 5 Mentoring Program is continuing at the 2025 APA Convention!

Overview of the Mentoring Program

As part of advancing practice and understanding of both quantitative and qualitative methodologies in psychology, Division 5 has hosted a mentoring program for the past three years. This program connects graduate students and early career scholars with more established academics and professionals who serve as mentors and provide guidance on specific methods and projects. Through this mentoring program, mentees can benefit from this guidance and collaboration, furthering their professional growth, while mentors gain experience working with early career scholars at different institutions and contribute to Division 5's mission. The goal is for these mentoring relationships to be meaningful, beneficial, and enduring. To that end, mentors and mentees are carefully paired together based on interests, goals, and experiences, and then provided with guidance to ensure the success of the mentor-mentee partnership.

What to Expect in the Mentoring Program

Mentors and mentees have been matched and they are set to communicate with each other and discuss the goals of their mentoring relationship. To facilitate this process, we will hold a workshop on Zoom in late July the APA convention for all mentors and mentees to attend. During the workshop, we will provide resources to support both mentors and mentees, discuss the goals and expectations of the mentoring program, and allow mentor-mentee pairs to get to know one another. By meeting as a large group, we hope that mentors and mentees will benefit not only from their partnership but also from connections with us and the other mentoring program participants.

Then, during the APA convention in Denver, we will hold an in-person session for mentors and mentees to meet up and discuss their projects. For those not attending the conference, there will be a virtual option. This meeting will allow participants to connect meaningfully with each other and build a support network of other division members. In addition, mentors and mentees are invited to join the Division 5 social at the conference (more details to come).

To recognize the program's participants who engage well with their mentor or mentee, they will receive an electronic certificate of participation signed by division leadership. We hope that this meaningful token of involvement will serve as a motivator for sustained participation and underscore the value we place on their contributions to the Division 5 community!

While we hope that mentors and mentees will continue to work together long after their involvement in the Division 5 mentoring program, the expectation is that they will work together for at least one year, meeting (virtually or in person) at least once a month. This duration and frequency will increase the likelihood of a successful and enduring partnership, while not placing an undue burden on either participant.

Previous Success and Participant Feedback

For the past three years, the mentoring program has successfully paired early career mentees with established mentors. Participants have provided positive feedback about their experiences in the mentoring program, particularly regarding professional development, enhanced understanding of advanced methodologies, and building and strengthening their professional networks. Mentors have valued the ability to contribute their knowledge and expertise to benefit emerging scholars, while mentees have valued the ability to gain personalized advice and guidance from experts in the field. We hope to have another successful mentoring program this year and will work closely with mentors and mentees to ensure they are supported at every step in the process.

Conclusion

The Division 5 Mentoring Program is designed to enhance the professional development of both emerging scholars and established professionals, and thus to contribute to the mission and strength of our division. Whether you are an emerging scholar seeking advice and support, or an established scholar who would like to provide guidance and expertise, our goal is that you will benefit from participating in this program.



Elizabeth Sanders

DIVISION 5 APA SESSIONS AND EVENTS

Elizabeth Sanders, PhD,

If you are attending APA, we cordially invite you to attend the following exciting Division 5 sessions:

Session	Type	Day	Time
Recent Methodological Developments in the Analysis of Intensive Longitudinal Data	Symposium	Thu, Aug 7	9:30 - 10:30 am
Teaching Qualitative Methods: On the Learner's Tension between Presuppositions and Bracketing	Critical Conversations	Thu, Aug 7	1:00 - 2:00 pm
Learning to Conduct an Integrative Mixed Method Meta-Analysis <i>(CE Credit)</i>	Skill Building	Thu, Aug 7	2:30 - 3:30 pm
Indigenous Methodologies For Assessment Research and Practices: Past, Present, Possible Futures	Critical Conversations	Thu, Aug 7	4:00 - 5:00 pm
Division 05 Business Meeting	Business Meeting	Fri, Aug 8	8:00 - 9:00 am
Division 05 Presidential Address + Division Awards	Division Presidential Address	Fri, Aug 8	11:30 - 12:30 pm
Structural Equation Models with Social Network Data	Skill Building	Fri, Aug 8	1:00 - 2:00 pm
New Twists in IRT Modeling : CAT and CDM Innovations	Symposium	Fri, Aug 8	4:00 - 5:00 pm
Assessment Community Breakfast + Student Awards (Divisions 5, 12, & 40)	Social Hour	Sat, Aug 9	8:00 - 9:00 am
Division 05 Poster Session I: Synergies of Methodological Research & Practice	Poster	Sat, Aug 9	1:30 - 2:30 pm
Division 05 Poster Session II: Psychometrics & Analytics - Next Gen Methods for the Modern Era	Poster	Sat, Aug 9	3:00 - 4:00 pm

Note: Virtual posters will be accessible online, throughout the conference and beyond.



PUBLICATON CORNER

*Compiled by
Gill Sitarenios*

Improving content validity evaluation of assessment instruments through formal content validity analysis

Traditional approaches to content validity often evaluate whether test items are relevant to a construct, but lack depth in analyzing item behavior or completeness. A new method called Formal Content Validity Analysis (FCVA) addresses this by combining structured classification matrices with interrater agreement metrics. An extension of this approach, Bayesian FCVA (B-FCVA), integrates Bayesian techniques to improve reliability. Both methods offer rich insights into content validity while remaining efficient for researchers and experts to use.

Spoto, A., Nucci, M., Prunetti, E., & Vicovaro, M. (2025). Improving content validity evaluation of assessment instruments through formal content validity analysis. *Psychological Methods*, 30(2), 203-222.
<https://doi.org/10.1037/met0000545>

Interpretable machine learning for psychological research: Opportunities and pitfalls

Machine learning tools are increasingly used in psychology, not just for prediction but also to understand variable relationships. However, many models—like random forests and neural networks—lack transparency. This paper explores interpretability techniques that clarify how such models make predictions. It highlights both the potential of these tools and the risks of misinterpretation, especially when variables are correlated or interact. Simulated and real-world examples are used to guide researchers in responsibly applying interpretability methods.

Henninger, M., Debelak, R., Rothacher, Y., & Strobl, C. (2025). Interpretable machine learning for psychological research: Opportunities and pitfalls. *Psychological Methods*, 30(2), 271-305.
<https://doi.org/10.1037/met0000560>

Missing data: An update on the state of the art

Since a landmark 2002 paper, methods for handling missing data in psychology have evolved significantly. This update reviews major advancements in theory and practice, including improved implementations of full information maximum likelihood, Bayesian estimation, and multiple imputation. It also covers emerging areas like factored regression and multilevel model handling. The article offers a snapshot of the current methodological and software landscape, with practical guidance for researchers facing missing data.

Enders, C. K. (2025). Missing data: An update on the state of the art. *Psychological Methods*, 30(2), 322–339.
<https://doi.org/10.1037/met0000563>

Is exploratory factor analysis always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory–exploratory continuum

With the growing array of factor analysis methods, researchers face challenges in selecting the most appropriate approach. This study compares several techniques across the exploratory–confirmatory spectrum—including CFA, EFA, BSEM, and hybrid methods like ECFA and BCFA—using simulations and real data. Results suggest that exploratory methods often yield more accurate results, particularly when models are slightly misspecified. The authors offer practical recommendations, including a decision flowchart, to guide method selection based on specific research conditions.

Nájera, P., Abad, F. J., & Sorrel, M. A. (2025). Is exploratory factor analysis always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory–exploratory continuum. *Psychological Methods*, 30(1), 16–39. <https://doi.org/10.1037/met0000579>

Using natural language processing and machine learning to replace human content coders

Manual content coding in psychology is time-consuming and prone to human error. This paper demonstrates how natural language processing (NLP) and machine learning can automate this process. The authors trained models on existing human-coded data to predict content codes effectively. A simulation study revealed how factors like sample size, label distribution, and coder reliability affect performance. The findings lead to concrete recommendations on when and how NLP can reliably replace human coders in psychological research.

Wang, Y., Tian, J., Yazar, Y., Ones, D. S., & Landers, R. N. (2024). Using natural language processing and machine learning to replace human content coders. *Psychological Methods*, 29(6), 1148–1163. <https://doi.org/10.1037/met0000518>

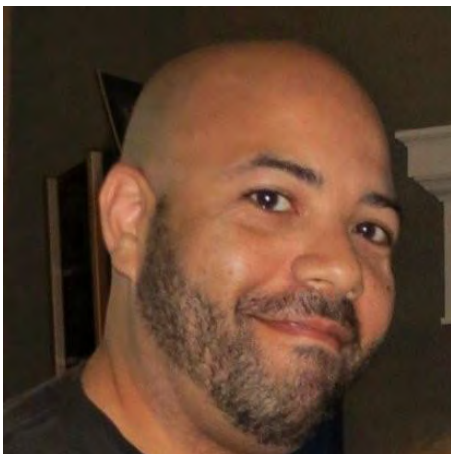


NEW SCORE EDITOR

After two terms and six years as editor of The Score, I am concluding my tenure as Editor of this newsletter. It has been a great privilege to work with Division 5 leadership who have unfailingly been kind and supportive. In particular, I'd like to thank the many presidents that have helped to oversee the newsletter during my terms: Rachel Fouladi, Fred Wertz, Deborah Bandalos, Alexander Beaujean, Eric Youngstrom, Kurt F. Geisinger, David Mackinnon, and Amy Schmidt. I'd also like to express my appreciation to my Assistant Editor, Joanna Solomon, who has been collaborating with me on this newsletter for the entire two terms.

I'm excited to announce that Kevin Williams, Ph.D., will be taking over as editor starting with the October, 2025 issue. A short summary for Kevin is given below. I know the newsletter is in good hands going forward and that Kevin will bring fresh ideas, excellent content and masterful organization to future issues.

Gill Sitarenios, Ph.D.



New Editor:

Kevin Williams, Ph.D.

Senior Measurement Scientist, ETS

Kevin Williams is the senior measurement scientist in the Center for Education and Career Development in the Research & Development division, joining ETS in 2016. He received his Ph.D. in personality psychology with a minor in quantitative psychology from the University of British Columbia in 2008. After coauthoring the seminal research article introducing the dark triad of personality, his research has expanded to include investigations into the predictive validity and malleability of

noncognitive constructs in workplace contexts, validity research for high school equivalency tests, psychometric evaluations of novel psychological assessments, law enforcement personnel selection, response bias in job performance evaluations, career and technical education (CTE) pathways, educational and career experiences of underrepresented groups, and identifying skills for the new economy through employer expectations. In addition to consistent activity as a journal article author and conference presenter, Williams frequently serves as an ad-hoc reviewer for a wide range of book chapters and journal manuscripts. Prior to joining ETS, his professional experience included the development and validation of psychological and licensure assessments for stakeholders in educational, workplace, clinical, and correctional fields.

LINKS TO RECENT PREVIOUS ISSUES OF “THE SCORE”



Issue Date	Vol #	Issue #	Link
April, 2025	47	2	https://www.apadivisions.org/division-5/publications/score/2025/04/issue.pdf
January, 2025	47	1	https://www.apadivisions.org/division-5/publications/score/2025/01/issue.pdf
October, 2024	46	4	https://www.apadivisions.org/division-5/publications/score/2024/10/issue.pdf
July, 2024	46	3	https://www.apadivisions.org/division-5/publications/score/2024/07/issue.pdf
April, 2024	46	2	https://www.apadivisions.org/division-5/publications/score/2024/04/issue.pdf
January, 2024	46	1	https://www.apadivisions.org/division-5/publications/score/2024/01/issue.pdf
October, 2023	45	4	https://www.apadivisions.org/division-5/publications/score/2023/10/issue.pdf
July, 2023	45	3	https://www.apadivisions.org/division-5/publications/score/2023/07/issue.pdf
April, 2023	45	2	https://www.apadivisions.org/division-5/publications/score/2023/04/issue.pdf
January, 2023	45	1	https://www.apadivisions.org/division-5/publications/score/2023/01/issue.pdf
October, 2022	44	4	https://www.apadivisions.org/division-5/publications/score/2022/10
July, 2022	44	3	https://www.apadivisions.org/division-5/publications/score/2022/07
April, 2022	44	2	https://www.apadivisions.org/division-5/publications/score/2022/04
January, 2022	44	1	https://www.apadivisions.org/division-5/publications/score/2022/01
October, 2021	43	4	https://www.apadivisions.org/division-5/publications/score/2021/10
July, 2021	43	3	https://www.apadivisions.org/division-5/publications/score/2021/07
April, 2021	43	2	https://www.apadivisions.org/division-5/publications/score/2021/04
January, 2021	43	1	https://www.apadivisions.org/division-5/publications/score/2021/01
October, 2020	42	4	https://www.apadivisions.org/division-5/publications/score/2020/10
July, 2020	42	3	https://www.apadivisions.org/division-5/publications/score/2020/07
April, 2020	42	2	https://www.apadivisions.org/division-5/publications/score/2020/04
January, 2020	42	1	https://www.apadivisions.org/division-5/publications/score/2020/01

Issues of The Score can be accessed at the following webpage:

<https://www.apadivisions.org/division-5/publications/score/issues>

DIVISION 5 MEMBERSHIP INFORMATION



Quantitative and Qualitative Methods

Division 5 draws together specialists in the areas of quantitative and qualitative methods. Membership includes specialists in the fields of educational psychology, industrial-organizational psychology, individual assessment, measurement, program evaluation, psychometrics, qualitative inquiry, research interviewing, research methods and statistics. Division 5 provides the opportunity to participate in this critical aspect of professional development, and is a vibrant community that serves the core interests and needs of rigorous scientific inquiry in the social sciences. Those who are interested in joining the division or who would like more information should be referred to the Division 5 website:

<https://www.apadivisions.org/division-5/membership>

Advertising in *The Score*

The Score is the newsletter of the American Psychological Association's Division 5—Evaluation, Measurement, and Statistics. Division 5 is concerned with promoting high standards in both research and practical application of psychological assessment, evaluation, measurement, and statistics. Approximately half of the Division 5 members are university faculty members in quantitative psychology, psychometrics, educational psychology, or industrial organizational psychology and half are engaged in careers in industry, including the areas of individual and large-scale assessment. More than 1,000 Division 5 members receive The Score each quarter. Advertisements in The Score may be in the form of display advertisements or job announcements. Both types of ads can include graphics and other design features and can be submitted as text or camera-ready display art. Prices for advertisements and size requirements are provided in the accompanying table. Submission deadlines are 45 days prior to publication: February 15 for the April issue, May 15 for July, August 15 for October, and November 15 for January.

To advertise in The Score, please contact Editor Gill Sitarenios at GSitarenios@yahoo.ca.

Size	Dimensions	Display Ad Price	Job Announcement Price
Full page	7.125" × 9.5"	\$235	Not available
Half page	7.125" × 4.75"	\$175	\$120
Third page	4.75" × 4.75"	\$125	\$90
	2.375" × 9.5"		
Sixth page	2.375" × 4.75"	\$90	Free/\$55
Notes: Insertion orders for four consecutive issues receive a 15% discount. First sixth page job ad each year free, thereafter \$55.			