

The Score



Quantitative and Qualitative Methods

DIVISION 5 PERIODICAL VOLUME 47 ISSUE 4

EDITOR: KEVIN M. WILLIAMS



In this issue . . .

BRUNO D. ZUMBO	2
Looking Through the Lens of Differential Item Functioning (DIF): Embracing the Many Ways of Being Human	
DANKA MARIC, GRANT A. FORE, BRANDON H. SORGE, FRANCESCA A. WILLIAMSON, & JULIA L. ANGSTMANN	6
What's in a Name? A Primer for Applying Content Analysis Principles in Factor Interpretation	
JOEL P. WIESEN	11
New Data on Behaviorally Anchored Rating Scales (BARS): Vanishing High Inter-Rater Reliability	
ALSO IN THIS ISSUE	15
Message from the Division 5 President (p. 15); In Memoriam: Dr. Arthur Ralph Hakstian (1941 - 2025) (p. 17); Congratulations to our 2025 Division 5 Award Winners (p. 18); Call for 2026 Editorial Fellowship Nominations (p. 21); Upcoming Inaugural Call for the Roger E. Kirk Division 5 Author of the Year Award (p. 22); Publication Corner (p. 24); Links to previous issues (p. 26); Division 5 Membership Information (p. 27); Advertising in The Score (p. 28)	

LOOKING THROUGH THE LENS OF DIFFERENTIAL ITEM FUNCTIONING (DIF): EMBRACING THE MANY WAYS OF BEING HUMAN



Bruno D. Zumbo, Ph.D.

Distinguished University Scholar & Professor
Tier 1 Canada Research Chair in Psychometrics and Measurement
University of British Columbia, Canada
bruno.zumbo@ubc.ca

How do we know if a test is truly fair? This article explores the evolution of differential item functioning (DIF)—from its early focus on detecting bias to today's fourth-generation perspective that embraces human diversity and the many ways people engage with tests.

Imagine two students—one from a rural school and one from an urban center—sitting for the same standardized test. Both have mastered the material, yet one struggles with certain items, not because of ability, but because the questions assume cultural references unfamiliar to them. Is the test really measuring what it claims to measure? This is the essence of differential item functioning, or DIF.

DIF may sound like a technical term buried in the language of psychometrics, but at its heart, it deals with a question everyone cares about: fairness in testing. Over the decades, the concept of DIF has grown from a narrow focus on biased test items into a broader framework for understanding how social, cultural, and psychological contexts shape test performance (Zumbo, 2007a, 2023).

This article highlights the key themes from a thread in my research program and connects them with the story I have come to call *From Rome to Vancouver*, tracing a journey from definitions and detection to explanation and inclusion.

Two Faces of DIF

DIF has always carried two meanings. First, it describes a phenomenon: when individuals from different groups—yet with the same underlying ability—respond differently to the same test item. Second, it refers to the statistical methods used to detect those differences.

The problem is that these two meanings often get blurred. Too many discussions of DIF jump straight to techniques, leaving its deeper purpose—ensuring fairness and valid interpretation of test scores—underexplored. Recognizing this duality helps us see DIF not only as a technical challenge but also as a way to think critically about the purpose of testing itself.

The Evolution of DIF Research

Over time, DIF research has unfolded in successive stages (Zumbo, 2007a):

- **First generation:** identifying item bias in high-stakes exams, often tied to legal and policy concerns about fairness.
- **Second generation:** clarifying the distinction between *impact* (true differences between groups) and *bias* (irrelevant item features) and refining statistical methods.

- **Third generation:** shifting the focus to explanation—asking why items behave differently and what this reveals about the test-taker's context, cognition, and culture.
- **Fourth generation (emerging):** reframing DIF as more than a problem to be fixed. Instead, it becomes a window into human diversity, offering insights into how people's experiences, histories, and communities shape the ways they approach tests (Zumbo, 2023).

This emerging fourth-generation perspective reflects a broader move toward explanation-focused validity: using DIF not just to flag items, but to understand human variation better and to design assessments that respect the richness of lived experience.

A Journey from Rome to Vancouver

The shift toward explanation took a major leap with the integration of DIF and explanatory modeling. On a long flight home from lecturing in Italy in 2003, the idea emerged: concepts of mediation and moderation—long used in psychology to study how and under what conditions variables affect outcomes—could illuminate DIF, and vice versa.

Soon, research showed how community and sociological factors could act as moderators or mediators of test performance (Zumbo & Gelin, 2005). For instance, gender roles, income level, or rurality may shape how students interpret and respond to items, even when their underlying ability is the same. Seen through this lens, DIF is not just about item wording or technical flaws. It is about how broader social contexts influence what happens when someone sits down to take a test (Zumbo et al., 2015; Zumbo, 2007a, 2023).

The Ecological Turn

My ecological model of item responding (Zumbo, 2007a, 2007b, 2023; Zumbo et al., 2015) is a conceptual framework that reimagines how we understand test-takers' responses to assessment items, converging on an ecological perspective—especially in the context of differential item functioning (DIF) and test validity.

Core Idea

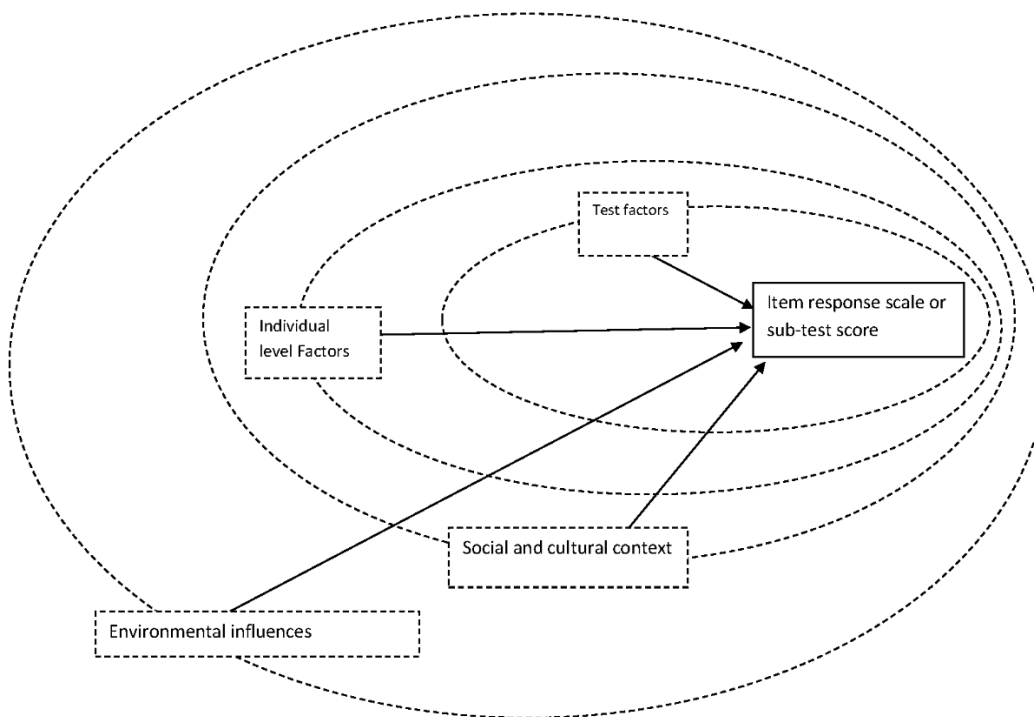
Rather than viewing item responses as isolated cognitive events, my model emphasizes that a complex ecology of interacting factors shapes responses. Instead of isolating the test-taker as an abstract individual, this approach situates them within layers of influence, for example, these may include:

- **Test factors:** Test format, item content, psychometric dimensionality
- **Individual-level factors:** cognitive abilities, language proficiency, motivation, prior knowledge
- **Social and cultural context:** norms, values, educational background, language of instruction
- **Environmental influences:** testing conditions, school infrastructure, community resources

In essence, the model argues that item responses are not just about the person or the item—they are about the *relationship* between the person and their environment.

Figure 1 is a graphical portrayal of an instance of my ecological model characterized by four concentric ovals described above.

Figure 1. A graphical portrayal of an instantiation of the ecological model characterized by four concentric ovals representing potential explanatory sources of variation for item, test, or scale scores, adapted from Zumbo et al. (2015), who consider an example of large-scale educational testing.



Depending on the in vivo assessment setting, different explanatory sources may be represented as concentric ovals. The outcome variable associated with these explanatory factors can take several forms: (i) responses to an individual item or task, (ii) a sub-scale score from the assessment or a specific dimension of a multidimensional measure, or (iii) an overall test score. Furthermore, depending on the chosen psychometric model, the sub-scale, dimension, or overall score may itself be conceptualized and modeled as a latent outcome variable.

This ecological view challenges the tendency to treat categories such as gender or ethnicity as fixed, "natural" kinds. Instead, they are better understood as social constructs—fluid, context-dependent, and shaped by history and culture (Zumbo et al., 2015; Zumbo, 2023; Zumbo & Lynch, in press). From this standpoint, DIF research becomes a way to study not only tests but also the environments in which they are used.

Why DIF Matters Today

In its modern form, DIF serves multiple purposes:

- Safeguarding fairness and equity across groups.
- Ensuring that observed differences are not measurement artifacts.
- Evaluating adapted or translated tests for comparability.
- Illuminating the cognitive and social processes of test-taking.
- Testing assumptions about invariance in advanced measurement models.

Rather than being just a statistical exercise, DIF helps explain how tests function in the real world. It also strengthens validity arguments by showing how scores arise, what they mean, and how they may differ across groups (Zumbo, 2007a, 2007b, 2009, 2023).

Bridging Methods and Meaning

Linking DIF with mediation and moderation also bridges psychometrics with broader psychological research. Both traditions remind us that statistical models are not neutral. They embed assumptions and generate inferences that shape how we understand test results.

This explanation-focused approach reframes validity not as a fixed property of a test, but as an ongoing, contextualized interpretation of score meaning. DIF, mediation, and moderation analyses all contribute to making those interpretations richer, more accurate, and more inclusive.

Looking Ahead

The journey from definitions of DIF to ecological and explanatory models marks a profound shift in measurement. DIF is no longer simply about detecting bias; it is about understanding the many ways human diversity shows up in testing.

The emerging fourth generation of DIF research reminds us that tests are not neutral tools and test-takers are not interchangeable units. Every testing event is shaped by the lived experiences, histories, and communities that individuals carry with them. Recognizing this, DIF becomes more than a technical safeguard: it becomes a lens for seeing the richness of human variation and the many ways of being human (Zumbo, 2023).

For psychologists, educators, and policymakers, this perspective underscores an enduring truth: tests are social occasions. By embracing ecological and explanatory approaches, DIF research can continue to advance fairness while helping us build a more inclusive and human-centered science of measurement.

Call to Action for Assessment Researchers and Psychometricians

As assessment researchers and psychometricians, we are uniquely positioned to carry this work forward. By incorporating fourth-generation thinking into our research, teaching, and applied measurement practices, we can deepen the field's commitment to equity and broaden its impact. The challenge before us is clear: to use DIF not only as a tool for ensuring fairness, but also as a framework for understanding and honoring the diverse ways people engage with tests.

Related Resources

- Zumbo, B. D., & Gelin, M. N. (2005). A matter of test bias in educational policy research: Bringing the context into picture by investigating sociological/community moderated (or mediated) test and item bias. *Journal of Educational Research & Policy Studies*, 5(1), 1–23. <https://eric.ed.gov/?id=EJ846827>
- Zumbo, B. D. (2007a). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly*, 4(2), 223–233. <https://doi.org/10.1080/15434300701375832>
- Zumbo, B. D. (2007b). Validity: Foundational issues and statistical methodology. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, Psychometrics, pp. 45–79). Elsevier. [https://doi.org/10.1016/S0169-7161\(06\)26003-6](https://doi.org/10.1016/S0169-7161(06)26003-6)
- Zumbo, B. D. (2009). Validity as contextualized and pragmatic explanation, and its implications for validation practice. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 65–82). Information Age Publishing.

- Zumbo, B. D., Liu, Y., Wu, A. D., Shear, B. R., Olvera Astivia, O. L., & Ark, T. K. (2015). A methodology for Zumbo's third generation DIF analyses and the ecology of item responding. *Language Assessment Quarterly*, 12(1), 136–151. <https://doi.org/10.1080/15434303.2014.972559>
- Zumbo, B. D. (2023). A dialectic on validity: Explanation-focused and the many ways of being human. *International Journal of Assessment Tools in Education*, 10(Special Issue), 1–96. <https://doi.org/10.21449/ijate.1406304>
- Zumbo, B. D., & Lynch, S. E. (in press). Intersectional psychometrics. In S. Sinharay (Ed.), *Encyclopedia of measurement in the social sciences*. Elsevier.

WHAT'S IN A NAME? A PRIMER FOR APPLYING CONTENT ANALYSIS IN FACTOR INTERPRETATION

Danka Maric¹, Grant A. Fore¹, Brandon H. Sorge², Francesca A. Williamson³, and Julia L. Angstmann⁴

¹STEM Education Innovation and Research Institute, Indiana University, Indianapolis

²O'Neill School of Public and Environmental Affairs, Indiana University, Indianapolis

³University of Michigan School of Medicine

⁴Office of Sustainability, Butler University



Danka Maric,
M.A.



Grant A. Fore,
Ph.D.



Brandon H.
Sorge, Ph.D.



Francesca A.
Williamson,
Ph.D.



Julia L.
Angstmann,
Ph.D.

Correspondence concerning this article should be addressed to Danka Maric, STEM Education Innovation and Research Institute, Indiana University, Indianapolis, 755 W. Michigan, UL1123, Indianapolis, 46202. Email: dmaric@iu.edu

In 1971, Paul Meehl and his colleagues describe a general uneasiness felt by many around factor analysis. They argue that the cause is the largely subjective nature of factor interpretation. That is, when it comes to factor interpretation and naming, it is “an after-the-fact, ad hoc enterprise” (Meehl et al., 1971). At the time, there were no satisfactory criteria for evaluating factor interpretation.

Today, measurement is a critical issue, and adequate psychometric evidence is imperative to obtaining valid and reliable research results (DeVellis, 2017). There is a plethora of resources guiding researchers on factor retention, loading strength, and rotation options (e.g., DeVellis, 2017; Flora & Flake, 2017; Tabachnick & Fidell, 2014; Yong & Pearce, 2013). Aside from the Recapture-Item Technique suggested by Meehl and

colleagues (1971), there is still a dearth of satisfactory guidance on factor interpretation. Indeed, factor naming is described as “more of an ‘art’” and as having no rules “except to give names that best represent the variables within the factors (Yong & Pearce, 2013).” Although we respect this method, we wish to offer an alternative approach for those that want something more systematic.

We suggest a mixed-methods approach, applying content analysis principles to the factor naming process. Content analysis is an analytical technique that involves detailed and systematic examination and interpretation of material with the goal of identifying patterns, themes, and meanings (Lune & Berg, 2017). It allows one to examine words, phrases, concepts, sentence structure, and semantics while considering theoretical frameworks and context. To illustrate, we describe our factor naming process for the Civic-minded Graduate (CMG; Steinberg et al., 2011) scale below.

Table 1

Items in Final Civic-minded Graduate Scale

Factor	Item
Interdependence of Self and Others	I want to dedicate my career to improving society.
	I would say that the main purpose of work is to improve society through my career.
	I intend to be involved in volunteer service after I graduate.
	I feel a deep conviction in my career goals to achieve purposes that are beyond my own self-interest.
	I believe that I have a responsibility to use the knowledge that I have gained to serve others.
	I have a sense of who I am, which includes a sincere desire to be of service to others.
Valuing Civic Duty	I believe that having an impact on community problems is within my reach.
	I intend to stay current with the local and national news.
	I stay up to date on the current political issues in the community.
Applying Interpersonal Knowledge	It is important for me to vote and be politically involved.
	I listen to others and understand their perspective on controversial issues.
	I am a good listener, even when people’s opinions are different from mine.
	Other students who know me well would describe me as a person who can discuss controversial social issues with civility and respect.
	When members of my group disagree on how to solve a problem, I like to try to build consensus.

Macro-awareness to Intervene	I am prepared to write a letter to the newspaper or community leaders about a community issue.
	I am aware of a number of community issues that need to be addressed.
	I am convinced that social problems are not too complex for me to help solve.
Micro-awareness to Intervene	I know a lot about opportunities to become involved in the community.
	I am able to plan or help implement an initiative that improves the community.
	I have the professional knowledge and skills that I need to help address community issues.
	I am very familiar with clubs and organizations that encourage and support community involvement for college students.

As part of a place-based experiential learning intervention study (Angstmann et al., 2019), we conducted a factor analysis of the CMG scale and found a 25-item, 5 factor structure (Maric et al., in press). Table 1 below displays the factors and items were retained in our structure. Initially, we (authors 1, 3, and 4) met and decided on preliminary factor names based on thematic interpretations and theoretical considerations. At this point, we decided that these names could be improved and wanted to be more systematic in our interpretations. Thus, using standard principles from content analysis (see Lune and Berg, 2017: 184), we examined items for sentence structure, coded them inductively, and then sorted them based on thematic categories.

We (authors 1, 2, and 3) independently considered each factor and their items and met to discuss our interpretations. Through deliberation, we each explained our reasoning behind our interpretations according to perceived patterns and theoretical considerations until consensus was reached. Theoretical frameworks included process philosophy (Whitehead, 1978), postcolonial theories of subjectivity (Nyamnjoh, 2017), and interdependent self-construal (see Markus and Kitayama, 1991). Table 2 below displays factor names, key words/phrases, and the themes that were extracted.

Table 2

Factor Names, Key Words/Phrases Identified, and Themes Extracted

Factor Name	Key Words/Phrases	Themes
Interdependence of Self and Others	"I"; "careers"; "purpose of work"; "conviction"; "responsibility"; "sincere desire"; "improving society"; "serv[ing] others"; "complex"	Connection to others and community; serving one's community
Valuing Civic Duty	"intend"; "vote"; "involved"; "stay"; "I" or "me"; "important"	Involvement with current events and politics

Applying Interpersonal Knowledge	“I”; “listen”; “discuss...with civility”; “build consensus”	Applying interpersonal skills such as consensus building and listening to others
Macro-Awareness to Intervene	“prepared”; “aware”; “convinced”; “I am”; “write”; “help solve”; “address”; issues”; “problems”	Knowledge of issues, engagement, and intervention relative to larger societal structures
Micro-Awareness to Intervene	“I”; “community”; “community issues”; “plan”; “encourage”; “support”; “involved”	Knowledge of issues, engagement, and intervention relative to local community structures

For all the factors, we paid close attention to sentence structure, including verbs and subjects. We noticed that in all the items, that the subject was the survey-taking student (denoted by the words “I” or “me”). We considered how these verbs related to the subject and what that meant. For example, in Factor 1, the items’ verbs articulate the intentionality, desire, feelings, sense of identity, and belief of the subject (“I”). In some items, one’s work, career, and specialized knowledge is directed toward others and society. The “I” is connecting their identity, desires, values, and work toward external relationships consolidated within the notions of society and community, leading to the name “Interdependence of Self and Others.”

In Factor 2, we noted that each question focused on intent to be either aware or involved in current local and national news and politics. Coupled with an expression of perceived importance in one item, we interpreted this as implying a valuing of some kind. We determined that these items were each articulating a relationship between one’s valuing of and intent to stay informed and engaged with larger structures of civil society (e.g., news and politics) and named this factor “Valuing Civic Duty”. In Factor 3, we found that there was significant emphasis on interpersonal relationality in diverse contexts, where unified solutions to complex or controversial issues were to be fashioned. Given that the items were all referring to interpersonal skills but also emphasized the application of those skills, we agreed upon “Applying Interpersonal Knowledge”.

The last two factors were admittedly difficult to differentiate. In Factor 4 the verbs were each preceded by an auxiliary verb articulating a state of being (“am”) that referred to a subject (“I”) – the student, connecting the person to an action. Each item also contained “issues” or “problems” in the predicates. Thus, these items were each articulating a relationship between one’s awareness and know-how relative to engagement with larger structures of civil society (e.g., newspapers, leaders), as well as problems and issues that are multiple and “complex.” Thus, we agreed upon “Macro-awareness to intervene. In Factor 5, the subject (“I”) was consistently in relation to the sentences’ objects, such as “community” and “community issues.” Like Factor 4, the verbs in these items referred to actions, this time taken in relation to community, which is on a more micro-level scale. Thus, we settled on the factor name: “Building Micro-Awareness to Intervene.”

Through this technique, we were able to align our factor names more closely with Dewey’s ethical and democratic theories (e.g., Dewey, 1976) underpinning the scales items. Previous iterations of the CMG’s structure (i.e., Steinberg et al., 2011) were not so explicitly aligned with these theories, despite their importance to the items. This also brought the CMG closer to recent conceptualizations of civic-minded education (Kreber, 2016). This paper serves as a primer for others to take a more systematic approach to interpreting and naming factors and a foundation for future work to build upon. Above all, this work can encourage researchers to engage in interdisciplinary collaboration, finding creative ways to incorporate and apply quantitative and qualitative methods into their work.

References

- DeVellis, R. F. (2017). *Scale development: Theory and applications*. Sage publications.
- Dewey, J. (1976). *The middle works, 1899-1924* (Vol. 13). SIU press.
- Kreber, C. (2016). *Educating for civic-mindedness: Nurturing authentic professional identities through transformative higher education*. Routledge.
- Lune, H., & Berg, B. L. (2017). *Qualitative research methods for the social sciences*. Pearson.
- Maric, D., Fore, G.A., Sorge, B.H., Williamson, F.A., & Angstmann, J.L. (in press). Validity and reliability of the civic-minded graduate scale in a place-based experiential learning context: Integrating ethical and self-construal theory. *The Good Society*.
- Markus, H. R., & Kitayama, S. (1991). Cultural and the Self: Implications for Cognition, Emotion, and Motivation. *Psychological Review*, 98(2), 224-253.
- Meehl, P. E., Lykken, D. T., Schofield, W., & Tellegen, A. (1971). Recaptured-Item Technique (RIT): A method for reducing somewhat the subjective element in factor naming. *Journal of Experimental Research in Personality*, 5(3), 171-190.
- Nyamnjoh, B. (2017). *Drinking from the Cosmic Gourd: how Amos Tutuola can change our minds*. African Books Collective.
- Steinberg, K. S., Hatcher, J. A., & Bringle, R. G. (2011). Civic-minded graduate: A north star. *Michigan Journal of Community Service Learning*, 17(1), 19-33.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6th ed.). Pearson.
- Whitehead, A. N. (1978). *Process and reality: Corrected edition*. NY: Macmillan.
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology*, 9(2), 79-94.

NEW DATA ON BEHAVIORALLY ANCHORED RATING SCALES (BARS): VANISHING HIGH INTER-RATER RELIABILITY



Joel P. Wiesen, Ph.D.

Applied Personnel Research
jwiesen@appliedpersonnelresearch.com

Abstract

We often misinterpret the high inter-rater reliability often obtained with behaviorally anchored rating scales (BARS). We commonly assume that high inter-rater reliability of grading panels indicates that scoring is reliable and, by inference, accurate. Rarely, if ever, do we evaluate the reliability between grading panels. Presented here are data from the scoring of 22 police promotional examination exercises. High within-panel Spearman-Brown grade reliability statistics sometimes coexist with low between-panel grade correlations. Between-panel correlations are rarely as high as would be expected based on the within-panel reliability of the pairs of panels. It may be that individual rater panels develop their own idiosyncratic grading criteria, despite all panelists receiving the same thorough training before live grading.

Introduction

Inter-rater reliability of grading with a BARS is generally high (e.g., 0.77, Levashina et al., 2014, page 273; and 0.84, McDaniel et al., 1994, Table 2 for structured interviews). Low inter-rater reliability is taken as an indication of grading deficiency and high inter-rater reliability is considered desirable (preferably 0.95 or higher, and at least 0.90, according to Nunnally & Bernstein, 1994, page 265). However, there is almost no literature that evaluates BARS rater- reliability across two rater panels scoring the same candidates on the same exercises. I encountered such data in my consulting work, as reported below.

Method

Data were available for 12 exercises administered as part of promotional examinations for police sergeant that were given biannually from 2013 to 2023. Each examination year included an oral panel component consisting of two new, independent exercises. Each exercise was graded by two independent rater panels, each composed of three raters. Different raters graded the two exercises. The raters were all current Sergeants or Lieutenants in comparable police departments located in various states other than the location of the examination. A small number of raters graded more than one examination. Each year there were somewhat more than 100 test takers. Some candidates took the promotional examination more than once, in different years.

The exercises simulated the work of a police sergeant including, for example, a one-on-one discussion with a subordinate experiencing some difficulties, addressing a group of subordinates on a critique of an emergency scene, or addressing a civilian or a group of civilians on topics such as community policing or crime in the neighborhood.

Each exercise was graded on four dimensions. The dimensions rated varied by exercise and year and typically included such dimensions as: Oral Communication, Interpersonal Relations, Supervision, Problem Analysis, and Problem Resolution. A candidate's score for an exercise was the average of 12 BARS ratings (3 raters on 4 nine-point BARS scales).

Rater training was conducted before any rating of actual test takers. This training included: job duties of a Sergeant, the BARS methodology, focusing on behaviors rather than general impressions, common rater errors, and practice in grading the exercises. In live rating, initial ratings were made prior to any discussion. Large discrepancies among the raters on any dimension were discussed followed by the raters making their final ratings. Consensus was not required. Only the final ratings are reported below but the pattern is very similar for the initial ratings.

Before COVID, from 2013 to 2019, one panel both administered and graded the exercise and the second panel graded the candidates based on video recordings. After COVID, in 2021 and 2023, the exercises were administered by video and both panels graded the candidates based on the video recording.

The Spearman-Brown reliability for the pairs of panels was used to predict the expected correlation between-panels using this formula (Ghiselli, Campbell, & Zedeck, 1981, page 242):

$$r_{xy} = \sqrt{(r_{xx})(r_{yy})}$$

Results

Within-panel Spearman-Brown reliability was high, averaging 0.91. The predicted correlation between-panels, based on within-panel reliability, was also high, averaging 0.91. However, the observed correlation between-panels was lower, at 0.60. The ratio of the observed to the expected between-panel correlations, expressed as a percent, averaged 66%, meaning there was a 34% average shrinkage from the expected to the observed between-panel correlation. (See Table 1.) Parenthetically I note that 8 exercises from 4 promotional examinations for Lieutenant, with from 24 to 38 test takers, show the same pattern of shrinkage for some but not all exercises.

Table 1. Within-panel Reliability and Between-panel Correlation by Year and Exercise

Year	Ex.	Reliability Within-panel 1	Reliability Within-panel 2	Observed Correlation Between-panels	Expected Correlation Between-panels	Percent (observed/expected)	Shrinkage
2013	1	0.92	0.92	0.46	0.92	50%	50%
2013	2	0.89	0.86	0.48	0.87	55%	45%
2015	1	0.89	0.95	0.48	0.92	52%	48%
2015	2	0.89	0.87	0.77	0.88	87%	13%
2017	1	0.88	0.80	0.76	0.84	90%	10%
2017	2	0.85	0.89	0.66	0.86	76%	24%
2019	1	0.97	0.91	0.44	0.94	47%	53%
2019	2	0.97	0.94	0.66	0.96	69%	31%
2021	1	0.86	0.90	0.48	0.88	55%	45%
2021	2	0.96	0.96	0.71	0.96	74%	26%

2023	1	0.94	0.94	0.64	0.94	68%	32%
2023	2	0.94	0.94	0.70	0.94	74%	26%
Means	0.91	0.91	0.60	0.91	66%	34%	

Discussion

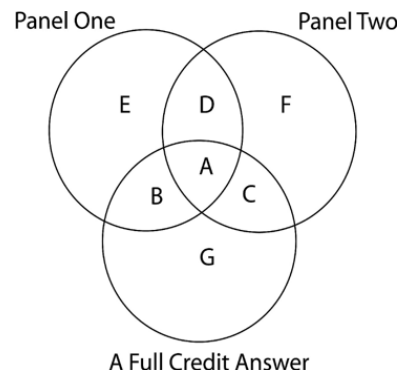
Possible explanations for these findings can be visualized with a Venn diagram (see Figure 1). On the positive side, the use of two rater panels may increase validity (see areas B and C, as each panel may capture valid aspects missed by the other panel). On the negative side, the observed agreement between panels may include invalid variance (see area D), due to scoring of invalid factors (e.g., a person's physical size). Areas E and F must include systematic as well as random error since the within-panel reliability is always 0.85 or higher. The systematic error not shared between panels may reflect idiosyncratic, panel-specific grading criteria (perhaps similar to measurement error).

The lower circle in Figure 1 is labeled “A full credit answer” rather than “The correct answer” because there may be more than one correct way to address complex work situations. The possibility of multiple correct answers could explain some of the shrinkage seen in Table 1, but is not considered further here.

One implication of this research is clear: using two BARS rating panels to grade complex job simulation exercises is now highly desirable, or even best practice.

Future research should try to identify the cause of low between-panel correlation seen in the grading of some exercises. Such future research might include:

- (1) having each pair of panels meet together to talk about the candidates whose scores show the most discrepancies,
- (2) having the agency staff who facilitated each pair of panels meet to talk about the candidates whose scores show the most discrepancies,
- (3) having the agency staff who facilitate the rating panels swap places midway through the grading to try to detect differences in grading criteria between panels,
- (4) doing a content analysis of the various exercises to try to identify the type of exercise that have low or high between panel correlations, and
- (5) analyzing the data to try to identify the dimensions associated with low or high between panel correlations.

Figure 1. Venn Diagram Showing Possible Valid and Invalid Variance in Grading

Key:

A: Valid variance identified by both rater panels

B: Valid variance identified by panel 1 but not panel 2

C: Valid variance identified by panel 2 but not panel 1

D: Invalid agreement between the two panels

E: Invalid variance exhibited only by panel 1 (both random and panel specific)

F: Invalid variance exhibited only by panel 2 (both random and panel specific)

G: Valid variance missed by both panels

Conflict of Interest Statement

I gained access to the examination statistics in the first 5 columns of Table 1 in my consulting role. These statistics were computed by the client as part of their standard post-examination analyses. The client did not influence my interpretation or the content of this article. I have no other potential conflicts of interest to disclose.

References

- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco, CA: Freeman.
- Levashina, J., Hartwell, C. J., Morgeson, F. P. & Campion, M. A. (2014) The structured employment interview: narrative and quantitative review of the research literature. *Personnel Psychology*, 67, 241- 293.
- McDaniel, M. A., Whetzel, D. A., Schmidt, F. L. & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-616.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.) New York: McGraw-Hill.

MESSAGE FROM THE APA DIVISION 5 PRESIDENT

Jolynn Pek, Ph.D.

Ohio State University
Pek.5@osu.edu



Background

It is both an honor and privilege to serve as the 2025-2026 President of Division 5. My journey with the Division began when I had the opportunity to support Debbi Bandalos (Past-Past-Past President) as Program Co-Chair for the APA Convention in 2023, alongside Wes Bonifay. That experience sparked a deep appreciation for what makes Division 5 truly special: the remarkable diversity of research interests, all connected by a shared commitment to advancing psychological methodology.

Past-Past Division 5 President Fred Wertz eloquently [addressed](#) the enduring challenge of balancing heterogeneity and unity within psychological methodology. Before him, Past-Past-Past-Past Division 5 President Alex Beaujean [chronicled](#) the evolving identity of Division 5, culminating in the formation of our three distinct sections: Evaluation, Measurement, and Statistics (EMS), Assessment, and Society for Qualitative Inquiry in Psychology (SQIP). Despite these structural changes, Alex emphasized a vital throughline – methodology – that continues to unify our Division’s work.

While *methods* often focus on tool development or the “what” used to forward psychological science, *methodology* delves into the “how” and “why” or the reasoning and justification behind the use and application of these tools. Prioritizing methodology over method is essential to enhance the rigor and impact of psychological science. I have written more on this perspective in the [July 2023 issue of *The Score*](#).

Initiatives

As President, I remain deeply committed to emphasizing methodology as a unifying force across the diverse perspectives of EMS, Assessment, and SQIP. This focus not only strengthens our collective identity but fosters meaningful collaboration across our Division’s sections.

One exciting initiative that exemplifies collaboration across our sections is the forthcoming *Handbook of Methodology for Psychology in the 21st Century: Progress, Critique, and Future* to be published by Guilford Press. Spearheaded by Past-Past President Fred Wertz, the project is being co-edited alongside former Division 5 Presidents Lisa Harlow and Rachel Fouladi. The comprehensive volume will span a wide range of methodological domains, including:

- Qualitative Inquiry – co-edited by Cynthia Winston and Lisa Osbeck
- Quantitative Analysis – co-edited by myself and former Division 5 President Debbi Bandalos
- Measurement and Assessment – co-edited by former Division 5 Presidents Kurt Geisinger and Alex Beaujean
- Evaluation – co-edited by Bianca Montrosse-Moorhead and Rodney Hopson
- Mixed Methods – contributed by Burke Johnson

This handbook sets the table by featuring the breadth and depth of methodology within our community. More than a reference, it serves as a gateway to methodology-minded scholars and practitioners to engage with

the rich diversity of approaches that shape our field. Each chapter not only traces the historical foundations of a given methodology but also envisions its future, offering a portrait of where we have been – and where we are headed.

Beyond intra-Division collaborations, Division 5 continues to collaborate with other Divisions in advancing methodological excellence. One example is the [Committee on Division/ APA relations \(CODAPAR\) interdivisional grant](#) proposal titled *Retaining History to Keep the Science and Practice of Measurement Alive*, which was recently submitted by Kurt Geisinger (President-Elect of Division 26, Society for the History of Psychology and former President of Division 5). This initiative brings together Divisions 5, 14 (Society for Industrial and Organizational Psychology), and 40 (Society for Clinical Neuropsychology), highlighting the enduring significance of measurement methodology across diverse psychological domains.

I look forward to continuing the tradition of collaboratively shaping Division 5 initiatives with the Presidential Trio – Past President Rachel Fouladi and President-Elect Leo Wilton – while also deepening communication and collaboration with the Division's Executive Committee. A central priority during my term is to foster meaningful member engagement by exploring how Division 5 can better support its community through both established programs, such as our Mentor Program, and new opportunities for networking and connection. To better understand the evolving needs and interests of our members, I am currently working with members of the Executive Committee to develop a membership survey. Your responses will help inform future initiatives and ensure our efforts are both responsive and inclusive. I warmly encourage you to participate in the survey and share your perspectives on how Division 5 can best serve you. Our initiatives are continually evolving, and your voice is essential in shaping their direction. Beyond the survey, I welcome your input at any time. Please feel free to reach out to me via [email](#) as I would love to hear your thoughts. Finally, the strength of Division 5 rests on the dedicated service of its members. If you are interested in contributing to the Division's leadership and initiatives, I invite you to connect with me about opportunities to serve.

2026 APA Convention Program

Looking forward to the 2026 APA Convention in San Francisco, I am thrilled to announce our program co-chairs: Zachary Fisher and Ken Kelly. Recognizing the transformative impact of Artificial Intelligence (AI) on psychological science, practice, and pedagogy, the 2026 convention theme is: *Psychological methodology meets artificial intelligence: Modern innovations and future advancements*.

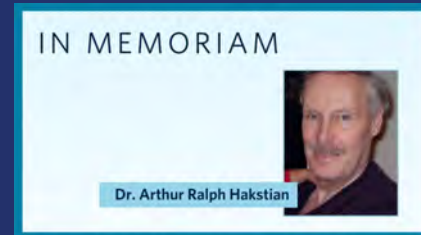
We welcome submissions exploring the transformative intersection of artificial intelligence and psychological science, particularly for advancing the field of methodology. As AI technologies reshape research landscapes, psychologists are uniquely positioned to advance both the science and practice of AI-enabled measurement, assessment, statistical analysis, and qualitative approaches, as well as pedagogy in methodology. Important conversations are also needed to address the logical and ethical limits of AI use in psychological science and research.

We invite presentations that demonstrate how AI can be used to complement or improve psychological research methods (quantitative, qualitative, and mixed methods), statistical approaches, and measurement science. We also invite presentations that examine the broader implications, best practices and ethical considerations of integrating AI into psychological research and practice.

As AI continues to shape the future of psychological methodology, our diverse membership – spanning disciplines, perspectives, and methodological expertise – is uniquely positioned to engage in meaningful dialogue about the opportunities and challenges AI presents. Through mentorship, networking, and the exchange of ideas, we aspire to foster a vibrant community where members can learn from one another and refine methodological

practices in response to emerging technologies. We look forward to your insights and contributions as we navigate this exciting frontier together.

IN MEMORIAM:
DR. ARTHUR RALPH HAKSTIAN
(1941 - 2025)



*Reprinted with permission from the University of British Columbia's Psychology Department website:
<https://psych.ubc.ca/news/in-memoriam-dr-hakstian/>*

On July 21, 2025, the Department of Psychology lost a valued member of our community, Professor Emeritus Arthur (Ralph) Hakstian.

It is with great sadness that we announce the passing of Dr. Ralph Hakstian, an integral and valued member of the UBC Psychology community. Dr. Hakstian distinguished himself in his scholarship, teaching, and practice as a quantitative psychologist, psychometrician and industrial/organizational psychologist. He completed his undergraduate training at UBC and his graduate studies at the University of Colorado.

Over his academic career, he held faculty positions at the University of Massachusetts, the University of Alberta, and the University of British Columbia, from which he retired with 27 official years of service. At UBC, he taught at the undergraduate and graduate level, worked and published with colleagues across different areas of the Department, and was instrumental in the development of graduate training programs in Psychometrics/Quantitative Psychology and in Industrial/Organizational Psychology.

Throughout his distinguished career, Dr. Hakstian was known not only for his rigorous and innovative research but also for his mentorship and generosity. He mentored his students with care and continued to publish and collaborate with them on consulting projects after their graduation. His commitment to academic excellence and intellectual curiosity shaped generations of students and researchers.

His scholarship included innovations in methods for multivariate statistics including correlation analysis, factor analysis, and scale development. Dedicated to the field of industrial/organizational psychology, his work also extended to areas of research in education, personality, forensic psychology, cognition, perception, and health behaviours.

We send our deepest condolences to Dr. Hakstian's family, friends, loved ones, former students, and colleagues.

CONGRATULATIONS TO OUR 2025 APA DIVISION 5 AWARD WINNERS!

Emily J. Shaw



Division 5 offers nine prestigious awards honoring outstanding contributions in the areas of quantitative, qualitative, and mixed methods in psychology. In August, we gathered at the APA Convention in Denver to honor the individuals who received our division's awards this year. As Chair of the Division 5 Awards Committee, it was a pleasure to publicly recognize their impressive work and to acknowledge their nominators, as well.

I'd like to thank the other members of the Division 5 Awards Committee, including Deborah Tolman, Donald Brown, Pani Farvid, Qiwei He, and Amy Elizabeth Schmidt. I appreciate the thoughtful care and consideration they brought to this work. Additional thanks goes to our previous Division 5 President, Rachel Fouladi, for her guidance and support throughout the Awards process.

Please join me in congratulating our 2025 Division 5 Award Winners and learning more about them.



The **Anne Anastasi Dissertation Award** recognizes distinguished dissertations that address a topic in the field of quantitative research methods (assessment, evaluation, measurement, statistics). This year's award goes to **Dr. Gyeongcheol Cho** at the Ohio State University.

Dr. Cho's dissertation, *Constructs May or May Not Be Latent: Studies on Two Domains of Structural Equation Modeling*, presents important methodological contributions to two domains of structural equation modelling (SEM): factor-based and component-based. This work offers researchers validated tools to better capture the complexities of human behavior and psychology within the SEM framework, allowing for more nuanced and effective use of SEM in psychology. His methods improve SEM's data-analytic capabilities and further its adoption by researchers and practitioners across various sciences as a main statistical methodology. It is also notable that he developed a free stand-alone SEM software program, based on his dissertation work, available to all.



The **Distinguished Dissertation in Qualitative Inquiry Award** recognizes distinguished dissertations that addressed a topic in the field of qualitative research methods. This year's award goes to **Dr. Leo Kiralla** at Bellevue College.

Dr. Kiralla's dissertation, *Recognition and The Embedded Self: Rhythmicity, Responsibility, and Paradox in Adult Monozygotic Twins*, studies how monozygotic twins navigate psychological and existential challenges in navigating the binary of individuation and maintaining connection. It is highly innovative in both its question and its methodology. In his nomination, Mark Freeman described Dr.

Kiralla's dissertation as "a potentially groundbreaking effort," including a literature review...that serves as a model for literature reviews in qualitative inquiry." Ruthellen Josselson, who chaired the dissertation, provided an additional letter, stating, "In its range, depth, insight, sophistication, creativity and writing, this is simply the best dissertation I have ever chaired."



The **Anne Anastasi Distinguished Early Career Contributions Award**

recognizes an early career individual who has made outstanding contributions to quantitative research methods. The contribution may be in academic, clinical, government, industry or other settings. It is sponsored by Fordham University in memory of Anne Anastasi, PhD. This year's award goes to **Dr. Jessica Kay Flake** at the University of British Columbia.

Dr. Flake is currently an Associate Professor in the Quantitative Methods at the University of British Columbia. Her research tends to focus on raising the bar on conducting high quality research, with an eye toward rigorous measurement practices and construct validity. The multidisciplinary nature of her work is also notable, connecting areas such as philosophy of science, cultural studies, psychology, and metascience, among others. Dr. Flake's work has been published in some of the top journals – both methodological and in psychology.



The **Distinguished Early Career Contributions in Qualitative Inquiry Award**

recognizes an early career individual who has made outstanding contributions to qualitative research methods. The contribution may be in academic, clinical, government, industry or other settings. This year's award goes to **Dr. Nisha Gupta** at the University of West Georgia.

Dr. Gupta's accomplishments in qualitative inquiry are quite notable, and she is among the most innovative qualitative researchers in psychology, having garnered a national and international reputation. Through her scholarship with South Asian queer women and theoretical writing on qualitative inquiry and critical epistemologies, she has, as her nominator Dr. Sunil Bhatia, observed "established creative arts as a legitimate qualitative method for understanding human experience, telling new stories, and reaching out to larger audiences." Dr. Gupta has published widely in high-impact psychology and qualitative inquiry journals and book chapters, and received numerous awards.



The **Jacob Cohen Distinguished Contributions to Teaching and Mentoring Award**

recognizes demonstrated excellence in teaching and mentoring within the field of quantitative research methods. This year's award goes to **Dr. David Budescu** of Fordham University.

Dr. Budescu's career spans nearly fifty years at primarily 3 top-ranked institutions. Most recently, he was the Inaugural Anne Anastasi Professor of Psychometric and Quantitative Psychology at Fordham from 2008 to 2024. Dr. Budescu is a highly recognized scholar in decision science and multivariate analysis. But this particular award is about his teaching and mentoring contributions to the field of quantitative research methods.

Dr. Budescu has been the primary mentor on 27 PhD dissertations and 17 master's theses. Dr. Budescu has mentored students into producing over 40 first-authored publications, with many more as contributing authors in peer-reviewed articles, conference proceedings and invited talks. Some of these student manuscripts have become quite influential in the areas of quantitative psychology, and his students received numerous awards for research they worked on under Dr. Budescu's supervision. His students have assumed positions in a variety of organizations, departments and universities, demonstrating his broad intellectual impact. The list of names, and their current roles, is impressive.



The **Distinguished Contributions to Teaching and Mentoring in Qualitative Inquiry Award** recognizes demonstrated excellence in teaching and mentoring within the field of qualitative inquiry. This year's award goes to **Dr. Sara McClelland** at the University of Michigan.

Dr. McClelland embodies excellence in teaching and mentoring undergraduate and graduate students, and postdoctoral fellows in qualitative inquiry and psychology overall. An award-winning teacher, she has consistently taught graduate seminars and modules in Qualitative Methods, Thematic Analysis, Mixed Methods, and

Interviewing in the Psychology Department. Her colleague Dr. Isis Settles noted that “her graduate courses are extremely popular with students in Psychology, as well as units across the university, which has made a once infrequent course, now a staple in the Psychology department’s curriculum. She has had only high teaching ratings scores across the board. Her students have described Dr. McClelland as “awesome, challenging, phenomenal, passionate, fair, respectful, intelligent, insightful, understandable, interesting, and intensely dedicated to each and every student.”



The **Distinguished Contributions in Qualitative Inquiry Award** honors an individual who has a long and distinguished history of scientific contributions within the field of qualitative research methods. This year's award goes to **Dr. Joseph P. Gone** at Harvard University.

Dr. Gone is considered the leading expert on Native American psychology among all psychologists and was the first recipient of the American Psychological Association's Henry Tombs Award, as well as the Stanley Sue Award (from APA's Division 12), catalyzing and advancing incredibly instructive, influential and deeply

important dialogues of Indigenous Research Methodologies in psychology, the social sciences, education, and Indigenous Studies. His research is characterized by consistent innovation in posing critical and urgent questions and Indigenous inquiry as necessarily collaborative and participatory, as well as the centrality of critical qualitative methods in Indigenous research methodologies.

Dr. Gone's work fuels collaborative programs that bring indigenous practices and knowledge to treatment of addictions and support of mental health to the communities that are affected. The significance of his work cannot be overstated, evidenced in his impressive productivity (including 116 peer-reviewed, 22 chapters,) and its extensive citation. He has been honored with more than 25 fellowships and career awards.



The newly established **Distinguished Contributions in Mixed Methods Award** honors an individual who has a long and distinguished history of scientific contributions to integrating qualitative and quantitative methods and in mixed methods methodology. This award is sponsored by Harris L. Friedman, PhD. This year's award goes to **Dr. John W. Creswell** at the University of Michigan.

Dr. Creswell is a professor of family medicine and Senior Research Scientist of the Michigan Mixed Methods Program at the University of Michigan. He has authored numerous articles and 33 books (including new editions) on mixed methods research, qualitative research, and research design.

Dr. Creswell's impact on the development and advancement of mixed methods is unparalleled. He is recognized as one of the co-founders of mixed methods approaches, with key contributions including his sole authored book, *Research design: Qualitative and quantitative approaches*, his co-edited books: *Handbook in mixed methods in social & behavioral research*, and *Designing and conducting mixed*

methods research; and as one of the co-founders and co-editors of the *Journal of Mixed Methods*, for which he now continues service on editorial board. He is also the founding President of the Mixed Methods International Research Association.

Note that in 2025, there were no nominations for the Division 5 Samuel J. Messick Distinguished Scientific Contributions Award.

Next Year's Awards

The deadline for the 2026 [Division 5 Awards](#) nominations will be on Monday, December 8th, 2025. Please mark your calendars, spread the word, and consider nominating a recent graduate, early career professional, or other well-deserving colleague.

CALL FOR 2026 EDITORIAL FELLOWSHIP NOMINATIONS

The following journals seek nominations for editorial fellowship positions. Fellows serve 12-month terms beginning January 1, 2026. Nomination deadlines are November 1, 2025 unless otherwise specified below. For more information and to submit applications visit <https://www.apa.org/pubs/journals/resources/call-editorial-fellowships>

American Psychologist	Developmental Psychology	Experimental and Clinical Psychopharmacology
Health Psychology	Journal of Applied Psychology	Journal of Comparative Psychology
Journal of Counseling Psychology	Journal of Experimental Psychology: Animal Learning and Cognition	Journal of Experimental Psychology: Applied
Journal of Experimental Psychology: General	Journal of Experimental Psychology: Learning, Memory, and Cognition	Journal of Family Psychology
Journal of Personality and Social Psychology: Attitudes and Social Cognition	Neuropsychology	Professional Psychology: Research and Practice (deadline: Nov. 15, 2025)
Psychological Assessment	Psychological Methods	Psychological Review
Psychology and Aging	Psychology of Addictive Behaviors	Rehabilitation Psychology

UPCOMING INAUGURAL CALL FOR THE ROGER E. KIRK DIVISION 5 AUTHOR OF THE YEAR AWARD

It is with pleasure that APA Division 5 would like to announce the upcoming call for the new Roger E. Kirk Division 5 Author of the Year Award!

The late Professor Roger E. Kirk (1930-2023) was a long-time member and Past-President of Division 5, distinguished author and dedicated teacher and Emeritus Distinguished Professor of Psychology and Statistics at Baylor University. Roger Kirk was committed to the pedagogy of quantitative methods and to Division 5. He provided a generous bequest with the American Psychological Foundation (APF) to establish the Roger E. Kirk Division 5 Fund, with the principal purpose of the Fund being to support specific annual programs in quantitative methods.

One of the programs to be funded is Roger E. Kirk Division 5 Author of the Year Award. The following provides details the terms of this specific award, including the anticipated Application Deadline in this inaugural cycle.

Later this fall, the APF will be publishing on its website (ampsyhfdn.org) the Request for Proposals for this award, along with information about the Related Program given to the home department of the Award recipient.

Program Name: Roger E. Kirk Division 5 Author of the Year Award

Application Deadline: January 15th, 2026

2026 budget award amount: \$50,000

Number of recipients: 1

Related program: Roger E. Kirk Division 5 Grant (\$15,000 given to the home department of the Award recipient)

Program Description: Below is a description of the program based on the fund agreement. The program description will be used for advertising and will be included on the program's APF web page.

The Roger E. Kirk Division 5 Author of the Year Award recognizes psychologists contributing to the body of work of quantitative methods. Preference is given to candidates who have written a book or document that could significantly improve the teaching of quantitative procedures or quantitative practices. Recipients will receive an award of \$50,000. In addition, the home department of the award recipient will receive an additional grant of \$15,000 to support quantitative and statistics activities, including research or teaching related to this area.

Program Elements:

Eligibility: APF encourages applicants from diverse backgrounds with respect to age, race, color, religion, creed, nationality, ability, sexual orientation, gender, and geography.

Applicants must:

- be psychologists with a doctoral degree at any stage
- demonstrate contributions to improving the teaching of quantitative procedures or quantitative practices

Self-nominations as well as nominations by others are encouraged for this award.

Preference will be given to:

- nominees who have written a book or document that could significantly improve the teaching of quantitative procedures or quantitative practices, which has been published by a scholarly press or in a peer review journal
- nominees who are full members or fellows of Division 5, for at least one year, at time of nomination
- nominees who are faculty members currently employed by a regionally accredited college/university or retired from such an institution

Application Materials:

- nomination statement that includes evidence of how the nominee's writings have contributed to improving the teaching of quantitative procedures or quantitative practices, making specific reference to the focal book/document authored by the nominee
- nomination statement will also state how the nominator will submit a copy of the focal book/document to the review committee upon request
- nominee's CV
- two additional external reviews/letters of support of the focal book/document authored by the nominee

Evaluation Criteria:*Nominees will be evaluated on:*

- impact
- innovation
- contributions to the field
- alignment with award's purpose

Supplemental Reading on Roger Kirk's legacy:

Beaujean, A. A., & Weaver, C. A. III. (2024). Roger Kirk (1930–2023). *American Psychologist*, 79(8), 1243.
<https://doi.org/10.1037/amp0001399>

Fogleman, L. (January 11, 2024). Baylor Mourns Death of Emeritus Distinguished Professor Roger E. Kirk.



PUBLICATION CORNER

Compiled by

Kevin M. Williams, Ph.D.

Qualitative Methodologies to Advance Anti-Racist Scholarship

Racism has persisted since the inception of what we now call the United States, causing immeasurable collective violence and trauma among Black, Indigenous, and People of Color (BIPOC). Although large-scale resistance among BIPOC communities have led to historical landmark legislation and policies, this is currently being threatened by the same White supremacist ideologies on which the United States was built. Using robust methods to capture the effects of racism in the here-and-now is crucial not only to bear witness to the violence and trauma it causes but also to better understand BIPOC communities' commitment to thrive, pursue justice, and experience joy and pleasure. Qualitative methods present rich approaches that embody the multidimensional experiences of marginalized members of society. It is our hope that this collection of articles document (1) the power of qualitative methods in understanding the impact of racism within BIPOC communities, (2) anti-racist frameworks that can be used by all qualitative researchers regardless of their method, and (3) ways in which BIPOC researchers are expanding qualitative methods to engage in anti-racist scholarship. As you read these articles, we urge you to see this special issue as the beginning of many conversations about how qualitative methods are not only useful but are robust approaches to document both the impact of racism on BIPOC communities and ways to be anti-racist scholars.

Abreu, R. L., & Williams, T. R. (2025). Qualitative methodologies to advance anti-racist scholarship: An introduction to the special issue. *Qualitative Psychology*, 12(2), 123–128. doi: 10.1037/qup0000336

Continuous Norming Approaches: A Systematic Review and Real Data Example

Norming of psychological tests is decisive for test score interpretation. However, conventional norming based on subgroups results either in biases or require very large samples to gather precise norms. Continuous norming methods, namely inferential, semi-parametric, and (simplified) parametric norming, propose to solve those issues. This article provides a systematic review of continuous norming. The review includes 121 publications with overall 189 studies. The main findings indicate that most studies used simplified parametric norming, not all studies considered essential distributional assumptions, and the evidence comparing different norming methods is inconclusive. In a real data example, using the standardization sample of the Need for Cognition-KIDS scale, we compared the precision of conventional, semi-parametric, and parametric norms. A hierarchy in terms of precision emerged with conventional norms being least precise, followed by semi-parametric norms, and parametric norms being most precise. We discuss these findings by comparing our findings and methods to previous studies.

Urban, J., Scherrer, V., Strobel, A., & Preckel, F. (2025). Continuous norming approaches: A systematic review and real data example. *Assessment*, 32(5), 654–674. doi: [10.1177/10731911241260545](https://doi.org/10.1177/10731911241260545)

Strengthening Human Epistemic Agency in the Symbiotic Learning Partnership with Generative Artificial Intelligence

This conceptual article explores shared epistemic agency between humans and generative artificial intelligence (GenAI), emphasizing the need to foster active human epistemic agency through adaptive epistemic stances. We propose a framework incorporating epistemic stances in interactions with GenAI, drawing on Tsai's (2004) work to explore the reinforcing relationship between them. The article outlines how learners' epistemic stances affect their interaction with GenAI and suggests strategies for fostering critical engagement and inquiry. We assert that learners with adaptive epistemic stances benefit more from GenAI and that continuous interaction with GenAI can reshape learners' epistemic development. This article aims to promote a symbiotic relationship where both human and machine contributions are valued, ultimately enhancing the overall learning experience through shared epistemic agency.

Wu, J.-Y., Lee, Y.-H., Chai, C. S., & Tsai, C.-C. (2025). Strengthening human epistemic agency in the symbiotic learning partnership with generative artificial intelligence. *Educational Researcher*, 54(6), 358-368. doi: [10.3102/0013189X251333628](https://doi.org/10.3102/0013189X251333628)

Research Methods in Indigenous Contexts

Research Methods in Indigenous Contexts is a qualitative-first, classroom-ready guide for work with Indigenous communities. It translates Indigenous epistemologies and ethics into steps for design, consent, analysis, and dissemination, with assignable activities, while indicating when to incorporate quantitative techniques—supporting mixed-methods instruction.

Groh, A. (2018). *Research methods in indigenous contexts*. New York City, NY: Springer. doi: 10.1007/978-3-319-72776-9

Comparison of Component-Based Structural Equation Modeling Methods in Testing Component Interaction Effects

Generalized structured component analysis (GSCA) and partial least squares path modeling (PLSPM) are statistical methods for component-based structural equation modeling, where constructs are represented as components or weighted sums of observed variables. While both methods employ different approaches to testing an interaction effect between components, their comparative performance in this context has not been systematically investigated. To address this gap, we conducted a simulation study to evaluate their relative performance across various experimental conditions. The results indicate that GSCA's approach consistently outperforms PLSPM's in testing a component interaction effect under all examined conditions.

Shen, Z., Cho, G., & Hwang, H. (2025). Comparison of component-based structural equation modeling methods in testing component interaction effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 32(5), 753–765. doi: [10.1080/10705511.2025.2497088](https://doi.org/10.1080/10705511.2025.2497088)

LINKS TO RECENT PREVIOUS ISSUES OF “THE SCORE”



Quantitative and Qualitative Methods

Issue Date	Vol #	Issue #	Link
July, 2025	47	3	https://www.apadivisions.org/division-5/publications/score/2025/07/issue.pdf
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



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 Kevin Williams at kmwilliams@ets.org.

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.			