

A Redesign of the ACGME Resident/Fellow Survey Through Expert Reviews and Cognitive Interviews

Anthony R. Artino Jr¹, PhD
 Kathy Malloy
 Rebecca S. Miller, MS
 Lynne M. Kirk², MD, MACP
 Timothy P. Brigham, MDiv, PhD

ABSTRACT

Background Since 2003, the Accreditation Council for Graduate Medical Education (ACGME) has surveyed residents and fellows in its accredited programs. The Resident/Fellow Survey is a screening and compliance tool that programs can also use for continuous quality improvement. However, stakeholders have reported potential problems with the survey's overall quality and credibility.

Objective To redesign the 2006 Resident/Fellow Survey using expert reviews and cognitive interviews.

Methods In 2018-2019, the ACGME redesigned the Resident/Fellow Survey using an iterative validation process: expert reviews (evidence based on content) and cognitive interviews (evidence based on response processes). Expert reviews were conducted by a survey design firm and an ACGME Task Force; cognitive interviews were conducted with a diverse set of 27 residents and fellows.

Results Expert reviews resulted in 21 new survey items to address the ACGME's updated accreditation requirements; these reviews also led to improvements that align the survey items with evidence-informed standards. After these changes were made, cognitive interviews resulted in additional revisions to sensitive items, item order, and response option order, labels, and conceptual spacing. In all, cognitive interviews led to 11 item deletions and other improvements designed to reduce response error.

Conclusions Expert reviews and cognitive interviews were used to redesign the Resident/Fellow Survey. The content of the redesigned survey aligns with the updated accreditation requirements and items are written in accordance with evidence-informed standards. Additionally, cognitive interviews resulted in revisions to the survey that seem to improve respondents' comprehension and willingness to respond to individual survey items.

Introduction

Since 2003, the Accreditation Council for Graduate Medical Education (ACGME) has surveyed residents and fellows in its accredited graduate medical education (GME) programs in the United States.^{1,2} The Resident/Fellow Survey serves 2 primary purposes: (1) the ACGME uses the survey results as a screening and compliance tool to assess residents' and fellows' perceptions of program quality; and (2) the survey provides individual programs and institutions with feedback that can be used locally for continuous quality improvement. The primary stakeholders for the Resident/Fellow Survey are the accreditors (ie, the ACGME Review Committees) and the GME programs' trainees, faculty, program directors, and institutional leadership.

In 2006, after 3 years of survey data collection, the ACGME reviewed stakeholder feedback and

consulted with a survey design firm to eliminate or rewrite poorly performing, confusing, or low variability items.² In 2018, the ACGME approved a major revision of the Common Program Requirements (CPRs) and accepted a new set of CPRs specific to fellowship programs, which became effective on July 1, 2019. Despite the improvements made to the Resident/Fellow Survey in 2006, as well as the validity evidence collected over the years,¹⁻³ stakeholders identified potential problems with the survey's overall quality and credibility.⁴⁻⁶

In light of the 2019 CPR revisions and continuing stakeholder concerns, the ACGME updated and redesigned the Resident/Fellow Survey in 2018-2019 using data collected from an iterative validation process. The purpose of this report is to describe this redesign effort, which was guided by expert reviews and cognitive interviews.

Methods

In the summer of 2018, the ACGME convened a task force comprised of 18 members: 3 ACGME Board members, 3 Review Committee chairs, 2 Review Committee resident members, 7 ACGME senior

DOI: <http://dx.doi.org/10.4300/JGME-D-22-00437.1>

Editor's Note: The online version of this article contains a summary of resident and fellow responses to several of the most challenging survey items, as well as their thoughts about those items and the respective response options.

leadership members, and 3 additional members (a facilitator, a public member, and a survey design consultant). Several of the task force members held or previously held key GME leadership roles: 1 was a medical school dean, 3 were designated institutional officials, and 9 were program directors. The task force's charge was to critically review the 2006 Resident/Fellow Survey and guide the redesign of the instrument. To accomplish this work, the task force wrote a request for proposals that resulted in submissions from 4 survey design firms. After review and in-person proposal presentations, the task force selected Research Triangle Institute (RTI) to conduct the survey redesign and initial validation in collaboration with the ACGME Task Force.

RTI and the task force employed Messick's unified theory of validity,⁷ as articulated in the *Standards for Educational and Psychological Testing*,⁸ to guide the survey redesign efforts using 2 sources of validity evidence: evidence based on content (ie, topics, wording, and format of individual survey items) and evidence based on response processes (ie, cognitive steps taken by respondents as they complete the survey and attempt to comprehend and respond to individual survey items).⁸ While most validity studies tend to emphasize other sources of validity evidence (eg, internal structure or relations to other variables),⁹ implicit in that work is the assumption that "respondents are able to understand the questions being asked, that questions are understood in the same way by all respondents, and that respondents are willing and able to answer such questions."¹⁰ RTI and the task force made no such assumptions in this validity study; instead, study efforts were concentrated on careful examination of the content of the 2006 Resident/Fellow Survey, and cognitive testing was employed to examine respondents' response processes in light of the changes made through expert reviews.¹¹ In short, the decision to examine these 2 sources of validity evidence was based on the need to first examine how respondents comprehend the individual survey items before moving on to psychometric testing.

Resident/Fellow Survey Background

Residents and fellows in ACGME-accredited programs complete the Resident/Fellow Survey annually to evaluate their perceptions and identify areas of compliance and potential non-compliance with program requirements.^{1,2} The 2006 survey instrument consisted of 46 predominantly close-ended, Likert-type items concerning 6 primary areas: educational content, available resources, adherence to work hour standards, evaluation participation, faculty supervision and teaching, and patient safety and teamwork.

Objectives

To redesign the 2006 Accreditation Council for Graduate Medical Education (ACGME) Resident/Fellow Survey using expert reviews and cognitive interviews.

Findings

Expert reviews and cognitive interviews resulted in a survey that better aligns with the ACGME's updated accreditation requirements, and employs evidence-informed survey items that respondents understand and are willing and able to answer.

Limitations

This study is limited by its use of content experts primarily employed by or affiliated with the ACGME, as well as by the small sample of volunteer residents and fellows interviewed.

Bottom Line

This work provides initial validity evidence for the redesigned Resident/Fellow Survey and its continued use as a tool for accreditation and continuous quality improvement.

The survey also assessed respondents' overall opinions of their GME programs.

Validation Procedures

Working with RTI, the task force redesigned the 2006 Resident/Fellow Survey using an iterative process that was conducted in 2 phases: (1) expert reviews of the 46-item Resident/Fellow Survey to identify problematic survey items and develop new items that addressed the updated CPRs (evidence based on content); and (2) cognitive testing of the revised Resident/Fellow Survey that resulted from the expert reviews (evidence based on response processes).⁸ To undertake these activities, the task force partnered with RTI over a 1-year period: the 2 groups met a total of 7 times between July 23, 2018 and September 23, 2019.

Phase 1: Evidence Based on Content: In phase 1, the task force reviewed the content of the 2006 Resident/Fellow Survey to assess its appropriateness, given the ACGME's updated CPRs. In addition, RTI's experienced survey methodologists examined the 2006 Resident/Fellow Survey using the Question Appraisal System.¹² The Question Appraisal System is a set of evidence-informed standards compiled from the literature; it includes a coding form that guides users through the systematic appraisal of survey items. The goal of this appraisal system is to help reviewers identify potential problems with the wording or structure of items that may lead to difficulties in survey administration, respondent comprehension, or other shortcomings.

Following their initial examination, RTI offered expert analysis of item quality and provided evidence-informed recommendations for how to improve the individual survey items to improve clarity and reduce

construct irrelevant variance.⁸ Their review focused on the following areas: (1) item order and ordering effects; (2) wording and number of items and response options; (3) response options that use ordinal quantifiers (eg, *sometimes* and *often*); (4) sensitive items like those about violations of work hour regulations; and (5) overall survey length.

Throughout phase 1, the 18 task force members served as the content experts and the RTI staff served as the technical survey design experts. RTI worked closely with the task force to identify problematic survey items, make recommendations, and write new items to address the updated CPRs. In all cases, the task force made final decisions regarding item revisions, additions, and deletions.

Phase 2: Evidence Based on Response Processes:

Cognitive testing, also referred to as cognitive interviewing, “is a psychologically oriented method for empirically studying the way in which individuals mentally process and respond to survey questionnaires.”¹³ As a qualitative method, cognitive testing relies on in-depth interviews designed to investigate whether a survey fulfills its intended purpose by collecting evidence based on cognitive processes.^{8,10,11,13}

Using the revised Resident/Fellow Survey that resulted from the expert reviews, RTI conducted 2 rounds of cognitive testing (between March 28 and June 6, 2019) to determine whether residents and fellows understood the intent of the revised survey items (and the overall survey) and could respond appropriately. In accordance with the cognitive testing literature,^{11,13} the sampling goal was to obtain input from respondents who are similar to the survey’s ultimate target population.

To get survey input from a diverse set of residents and fellows, RTI developed an 11-question web screener, which determined eligibility and collected contact information for interested residents and fellows. In their weekly e-Communication with GME programs, the ACGME published a link to the web screener on March 13, 2019. After several weeks, the web screener drew a total of 3197 resident/fellow volunteers, with unique identification numbers. Quotas were then determined using the proportion of each category from the web screener responses. Residents comprised 82% of the total responses, and fellows comprised 18%. For round 1, this equated to a goal of 12 cognitive interviews with residents and 3 with fellows. Quotas by year of residency were determined using the same procedure. Of the total number of responses, year 1 comprised 26%, year 2 comprised 26%, year 3 comprised 25%, year 4 comprised 13%, year 5 comprised 6%, year 6

comprised 2%, and year 7 comprised 1%. With years 6 and 7 representing the smallest population of responses, this totaled less than 1 interview of the total 15 interviews targeted for round 1. RTI used a similar procedure to recruit round 2 participants proportionate to the pool of volunteer responses, and all study participants were remunerated \$200.

Seven different RTI staff members were trained on the interview guide and conducted the cognitive interviews. Training included a review of the process for conducting the interview, how to take notes during the interview, and a review of questions of interest to the task force.

Interviewers conducted the cognitive interviews via web teleconference for up to 2 hours each. Interviewers shared the survey instrument (as a Word document) with respondents on screen, and respondents were asked to read each survey item and state their response out loud to the interviewer. Interviewers used an interview guide with scripted probes for questions of interest but asked ad-hoc probes that naturally arose during the interview. The interview guide also explored whether there were any topics missing from the survey. Interviewers recorded each cognitive interview and also took notes during the interview.

Changes to the survey that arose from round 1 of the cognitive interviews were incorporated into the survey and tested in round 2. During round 2 cognitive interviews, residents and fellows reviewed a programmed web survey, as opposed to reviewing survey items in a Word document. Reviewing the web survey allowed residents to experience the survey in a format similar to how it would ultimately be administered by the ACGME.

RTI reviewed cognitive interview data from rounds 1 and 2, summarized all resident/fellow comments, and then met with the task force to discuss the results. The task force approved all item revisions that resulted from the cognitive testing.

Data Collection and Ethical Review

RTI collected data during all phases of this work. The data included changes made to the survey during the expert reviews, as well as results from 2 rounds of cognitive testing.

This study was approved by the RTI’s Institutional Review Board.

Results

Phase 1: Expert Reviews and New Items

To address the updated CPRs, the task force wrote 21 new survey items about several topic areas, including diversity and harassment concerns, protected time for

TABLE 1

Summary of Item-Level Recommendations From the Phase 1 Expert Reviews

Identified Problem With Survey Item	Recommended Revision and Rationale
Sensitive wording: “How often did you break the rule that clinical experience and educational work hours must be limited to no more than 80 hours per week...”	The phrase “break the rule” implies intentional, deliberate wrongdoing. However, respondents may have unintentionally worked more hours than they were supposed to. Alternately, respondents may have been pressured or forced into working more hours than they wanted. As a result, respondents may be reluctant to admit to the behavior when worded this way. <i>Recommendation:</i> Revise the wording to ask how often the behavior occurred without referencing breaking rules or meeting requirements. “How often did you work more than 80 hours per week...”
Item order: The survey begins with very challenging questions about work hour violations.	<i>Recommendation:</i> Build rapport with respondents by beginning the survey with a few easier, nonconfrontational questions, and then ask about work hour compliance.
Response option order: Some survey item response options go from low to high (eg, <i>not at all to a great deal</i>), whereas other response options go from high to low (eg, <i>extremely satisfied to not at all satisfied</i>).	<i>Recommendation:</i> Order all response options in the same direction, preferably from low to high, which aligns with the manner in which respondents read (from left to right and top to bottom).
Response option labels and spacing: The response option labels used for some of the items are not equal distances apart, conceptually. For example, the labels <i>very</i> and <i>extremely</i> , as well as <i>somewhat</i> and <i>slightly</i> may be interpreted by respondents to be quite similar in the following set of response options: <ul style="list-style-type: none"> ▪ Extremely satisfied ▪ Very satisfied ▪ Somewhat satisfied ▪ Slightly satisfied ▪ Not at all satisfied 	<i>Recommendation:</i> Provide a balanced set of response options by ensuring that all of the verbal labels are relatively equal distances apart, conceptually. <ul style="list-style-type: none"> ▪ Not at all satisfied ▪ Slightly satisfied ▪ Moderately satisfied ▪ Quite satisfied ▪ Very satisfied

educational activities for residents and fellows, individualized learning plans for residents and fellows, responsibility for reporting patient safety events, and participation in interprofessional clinical patient safety activities and interprofessional quality improvement activities. In addition, during phase 1, RTI shared the findings from their methodological review employing the Question Appraisal System; these results are summarized in TABLE 1, along with a rationale for each recommended revision.

Phase 2: Cognitive Testing

A total of 27 residents and fellows (17 in round 1; 10 in round 2) participated in 2 rounds of cognitive testing using the Resident/Fellow Survey that was revised in the phase 1 expert reviews. Descriptive statistics for those interviewed are provided in TABLE 2, and below is an overview of several of the topics that emerged from the cognitive interviews.

Overall Reaction to the Survey: As a group, residents and fellows viewed the tested versions of the survey in

a positive way. Broadly speaking, participants reported that they thought the survey was clearly written and comprehensive. They noted that they understood the purpose of the survey and thought the bulk of the items were relevant and important (see exceptions below).

Participants recognized that some survey items had answers that would potentially violate ACGME rules (eg, violations of work hour standards), and that answers could have implications for a residency program. Residents and fellows were largely interested in avoiding adverse actions for their programs. While residents and fellows noted that they understood that the survey was anonymous, and they indicated that this anonymity was important to their decision to share details of their program experiences, some interviewees expressed hesitation about providing information that might demonstrate to their program that they were personally in violation of rules (even though individual trainee responses are not shared with programs).

When an interviewer asked a particular resident whether he was comfortable sharing details of his work hours, he responded,

TABLE 2
Resident and Fellow Cognitive Interview Participant Profiles (n=27)

Round 1		
Interview Date	Role	Specialty
3/28/2019	Resident	General pediatrics
3/28/2019	Resident	Anesthesiology
3/28/2019	Resident	Physical medicine and rehabilitation
3/28/2019	Resident	General surgery
3/29/2019	Resident	Ophthalmology
3/29/2019	Resident	Emergency medicine
3/29/2019	Resident	Family medicine
3/29/2019	Resident	Internal medicine
3/31/2019	Resident	Pathology
4/1/2019	Resident	Psychiatry
4/1/2019	Resident	Orthopedics
4/1/2019	Resident	Thoracic surgery
4/2/2019	Resident	Obstetrics and gynecology
4/3/2019	Resident	Obstetrics and gynecology
4/3/2019	Fellow	Adolescent medicine
4/3/2019	Resident	Pediatrics
4/4/2019	Resident	Emergency medicine
Round 2		
Interview Date	Role	Specialty
5/23/2019	Resident	Radiology
5/23/2019	Resident	Neurological surgery
5/24/2019	Resident	Pathology
5/28/2019	Resident	Obstetrics and gynecology
5/29/2019	Fellow	Pediatric hematology/oncology
5/30/2019	Resident	Urology
5/30/2019	Resident	Pediatrics, psychiatry, and child psychiatry
6/6/2019	Resident	Preventive medicine
6/6/2019	Fellow	Pediatric hematology/oncology
6/6/2019	Fellow	Cardiology

“(It is) always a little uncomfortable because you want your program to do well because you’re graded compared to other programs. In the back of your mind, you want your program to do well. And there’s another part of you where your pride comes in and you want to shine and show you’re working a lot and really hard. But if you really look at the time and numbers, most of the time your hours are normal. As a second year, working a lot and feeling the burn, it would’ve been a lot more uncomfortable to answer these” (second-year resident, round 1).

Application Across Specialties: Participants noted several places in the survey where items did not seem to apply to their specialty. For example, some residents and fellows noted that other specialties likely work more hours or have more pressure to exceed work hours than their own specialty. A pathology resident noted that her limited work with patients or on rounds affected which items were applicable to her. Several surgeons and anesthesia residents noted that they often needed supervision in their duties regardless of their ability, because of the nature of their work. An ophthalmology resident noted that she did not participate in rounds like residents from other specialties might. In light of these findings, RTI added “don’t know” options on several items to allow respondents to report that they are not aware or do not know the answers to specific questions.

Hover-Over Definitions: When reviewing the programmed web survey in round 2 of the cognitive interviews, several respondents commented on the “hover-over” definitions (ie, definitions that become visible on the screen only when the cursor is positioned over specific text). In several instances, residents and fellows reported that the hover-over definitions were problematic. Although several participants said the definitions were helpful, they also noted that they did not always choose to access them or did not know that the display of the hover-over definitions implied that there was a definition available. In addition, the hover-over definitions were found to be less effective when compared to including the definitions in the survey item itself. Based on these findings, RTI recommended that the hover-over definitions be removed in favor of adding definitions to the items themselves, when appropriate.

Item-by-Item Findings: The online supplementary data summarizes resident and fellow responses to several of the most challenging survey items, as well as their thoughts about those items and the respective response options. The supplementary table also summarizes RTI’s recommended edits to the items based on this cognitive interview feedback. Altogether, the cognitive interviews led to 11 item deletions.

Discussion

This study of the 2006 Resident/Fellow Survey employed expert reviews (evidence based on content) and cognitive interviews (evidence based on response processes) to uncover numerous areas for improvement. The iterative, multiphase validation process resulted in deletions and additions of survey items,

and changes to item wording and format throughout. Overall, the revised version of the survey was received positively by most of the residents and fellows engaged in cognitive testing.

In phase 1 of the validation process, new items were added to the survey instrument to align it with the updated CPRs. This is especially important because, since 2003, an ACGME guiding principle is to base Resident/Fellow Survey items on the CPRs. In addition, although some of the changes based on the Question Appraisal System and cognitive interviews seem minor, decades of research in cognitive psychology and public opinion polling¹⁴⁻¹⁷ suggest that survey instrument features, such as the wording of items and response options, item format, and item order, largely determine how respondents interpret and respond to a survey. Ultimately, studies show that poorly designed survey items negatively affect the quality of survey data and result in construct irrelevant variance. Overall, our study results suggest that the revisions made to the 2006 Resident/Fellow Survey are an enhancement and result in a more credible measurement tool.

Results from the cognitive testing underscore the challenge of creating a survey for a diverse group of residents and fellows who have a variety of responsibilities and experiences, as well as the need to carefully review and pretest each survey item. One area closely examined during both expert reviews and cognitive interviews was the use of so-called vague quantifiers to measure the frequency of perceived work hour violations (eg, *sometimes*, *often*). Vague quantifiers measure respondents' perceptions of frequency, not actual frequency. Several vague quantifiers were examined during the expert reviews and cognitive testing, and residents and fellows responded most consistently to the options tested in the final version: *never*, *almost never*, *sometimes*, *often*, *always*. Furthermore, a contemporary examination of vague quantifiers suggests they may have better measurement properties than numeric responses of behavioral frequency.¹⁸ Nonetheless, education scholars have criticized the ACGME's use of vague quantifiers.¹⁹ Therefore, follow-up psychometric analyses are needed to examine the distribution of the response scores, which could further corroborate (or refute) the initial response processes validity evidence reported here.

These study findings are limited by our use of content experts primarily employed by or affiliated with the ACGME, with few study participants acting as current program directors or designated institutional officials. ACGME employees are likely more familiar with CPRs, which enhances their expertise in content evaluation. But they may not have an

accurate sense of the key, current problems with the Resident/Fellow Survey from a user's perspective. In addition, despite attempts to recruit a diverse set of residents and fellows for interviews, the participants were volunteers who may have been more interested in or familiar with the Resident/Fellow Survey than non-volunteers. Finally, bias may have entered into our interpretations of the findings from the expert reviews and cognitive interviews, as final survey item decisions were made by task force members, who may have more favorable biases toward certain issues due to their ACGME affiliations.

Following this initial work examining content and response processes validity evidence, key next steps would include an examination of the survey's psychometric properties and internal structure using factor analysis and other statistical techniques. In addition, involving current stakeholders, such as program directors, teaching faculty, and designated institutional officials, is an important future consideration. Qualitative studies involving these individuals and other key stakeholders (eg, typical residents, fellows, and ACGME accreditation personnel) should be conducted to examine the usefulness of survey results for accreditation and continuous quality improvement activities.²⁰

Conclusions

Expert reviews and cognitive testing resulted in numerous changes to the 2006 ACGME Resident/Fellow Survey to align it with the updated CPRs and evidence-informed best practices in item writing. The cognitive interviews suggest that residents and fellows understand the revised survey questions and are willing and able to answer those questions. These results provide initial support for the quality and credibility of the redesigned Resident/Fellow Survey and its continued use as a tool for accreditation and continuous quality improvement.

References

1. Holt KD, Miller RS. The ACGME resident survey aggregate reports: an analysis and assessment of overall program compliance. *J Grad Med Educ*. 2009;1(2):327-333. doi:10.4300/JGME-D-09-00062.1
2. Holt KD, Miller RS, Philibert I, Heard JK, Nasca TJ. Residents' perspectives on the learning environment: data from the accreditation council for graduate medical education resident survey. *Acad Med*. 2010;85(3):512-518. doi:10.1097/ACM.0B013E3181CCC1DB
3. Holt KD, Miller RS, Vasilius J, et al. Relationships between the ACGME resident and faculty surveys and

- program pass rates on the ABIM internal medicine certification examination. *Acad Med*. 2018;93(8):1205-1211. doi:10.1097/ACM.0000000000002228
4. Colbert CY, Brateanu A, Nowacki AS, Prelosky-Leeson A, French JC. An examination of resident perspectives on survey participation and methodology: implications for educational practice and research. *J Grad Med Educ*. 2021;13(3):390-403. doi:10.4300/JGME-D-20-01431.1
 5. Fahy BN, Todd SR, Paukert JL, Johnson ML, Bass BL. How accurate is the Accreditation Council for Graduate Medical Education (ACGME) resident survey? Comparison between ACGME and in-house GME survey. *J Surg Educ*. 2010;67(6):387-392. doi:10.1016/J.JSURG.2010.06.003
 6. Sticca RP, MacGregor JM, Szlabick RE. Is the Accreditation Council for Graduate Medical Education (ACGME) resident/fellow survey a valid tool to assess general surgery residency programs compliance with work hours regulations? *J Surg Educ*. 2010;67(6):406-411. doi:10.1016/J.JSURG.2010.09.007
 7. Messick S. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *Am Psychol*. 1995;50:741-749.
 8. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association; 2014.
 9. Lyons-Thomas J, Liu Y, Zumbo BD. Validation Practices in the Social, Behavioral, and Health Sciences: A Synthesis of Syntheses. In: Zumbo BD, Chan EKH, eds. *Validity and Validation in Social, Behavioral, and Health Sciences*. Springer; 2014:313-319.
 10. Collins D. Pretesting survey instruments: an overview of cognitive methods. *Qual Life Res*. 2003;12(3):229-238. doi:10.1023/A:1023254226592
 11. Willis GB, Artino AR. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *J Grad Med Educ*. 2013;5(3):353-356. doi:10.4300/JGME-D-13-00154.1
 12. Willis G, Lessler J. Question Appraisal System QAS-99. January 1999. http://www.websm.org/db/12/15804/Web_Survey_Bibliography/Question_Appraisal_System_QAS99/
 13. Willis G. Cognitive Aspects of Survey Methodology. In: Lavrakas P, ed. *Encyclopedia of Survey Research Methods*. Vol 2. Sage; 2009:103-106.
 14. Dillman DA, Smyth JD, Christian LM. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. John Wiley & Sons Inc; 2014.
 15. Krosnick JA, Presser S. Question and Questionnaire Design. In: Marsden PV, Wright JD, eds. *Handbook of Survey Research*. 2nd ed. Emerald Group Publishing; 2010.
 16. Phillips AW, Artino AR Jr. Lies, damned lies, and surveys. *J Grad Med Educ*. 2017;9(6):677-679. doi:10.4300/JGME-D-17-00698.1
 17. Schwarz N. Self-reports: how the questions shape the answers. *Am Psychol*. 1999;54(2):93-105. doi:10.1037/0003-066X.54.2.93
 18. Baghal T Al. Is vague valid? The comparative predictive validity of vague quantifiers and numeric response options. *Surv Res Methods*. 2014;8(3 SE-Articles). doi:10.18148/srm/2014.v8i3.5813
 19. Yock Y, Lim I, Lim YH, Lim WS, Chew N, Archuleta S. Sometimes means some of the time: residents' overlapping responses to vague quantifiers on the ACGME-I resident survey. *J Grad Med Educ*. 2017;9(6):735-740. doi:10.4300/JGME-D-17-00187.1
 20. Caniano DA, Martinez SA, Nace C, Hogan SO. Reasons for data-prompted site visits: field staff findings and review committee decisions. *J Grad Med Educ*. 2021;13(3):447-454. doi:10.4300/JGME-D-21-00435.1



Anthony R. Artino Jr, PhD, is Professor and Associate Dean for Evaluation and Educational Research, the George Washington University School of Medicine and Health Sciences, and Deputy Editor, *Journal of Graduate Medical Education*; **Kathy Malloy**, is Vice President of Accreditation Standards, Accreditation Council for Graduate Medical Education (ACGME); **Rebecca S. Miller, MS**, is Senior Vice President of Applications and Data Analysis, ACGME; **Lynne M. Kirk, MD, MACP**, is Chief Accreditation Officer, ACGME; and **Timothy P. Brigham, MDiv, PhD**, is Chief of Staff and Chief Education Officer, ACGME.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

Disclaimer: A.R.A. was a paid ACGME consultant on this work. In addition, K.M., R.S.M., L.M.K., and T.P.B. are all full-time employees of the ACGME.

Corresponding author: Anthony R. Artino Jr, PhD, the George Washington University School of Medicine and Health Sciences, aartino@email.gwu.edu, Twitter @mededdoc

Received June 2, 2022; revisions received July 21, 2022, and November 26, 2022; accepted December 5, 2022.