Longwood University, Farmville, Virginia

. . . .

Authors

JOHN ERIC M. LINGAT
Assessment Coordinator,
Office of Strategic Planning
and Institutional Effectiveness
Department of Educational,
School, and Counseling
Psychology
University of Kentucky
johneric@uky.edu

XIAO-YIN CHEN
Graduate Research Assistant
Department of Educational
Psychology
University of Georgia
xy.chen@uga.edu

MICHAEL D. TOLAND *
Executive Director,
The Herb Innovation Center
University of Toledo
Michael.Toland@utoledo.edu

JAEYUN HAN
Graduate Research Assistant
Department of Educational,
School, and Counseling
Psychology
University of Kentucky
jaeyun.han@uky.edu

CHEN QIU
Graduate Research Assistant
Quantitative and Psychometric
Methods Program
chenqiu88@uky.edu

LIJUN SHEN
Graduate Research Assistant
Department of Educational
Policy Studies and Evaluation
University of Kentucky
lijun.shen.0509@uky.edu

Optimizing the Measurement of Resident Advisors' Self-Efficacy Using the Rasch Model

RESIDENT ADVISORS PLAY AN IMPORTANT ROLE IN STUDENT SUCCESS by providing peer leadership in residence halls. However, limited information and instruments are available to measure RA perceptions about their abilities to fulfill this role. This study proposes a new instrument and process to optimize the measurement of RA self-efficacy using locally developed items sourced from RA duties established by the host institution. The Rasch model was used to provide validity evidence for inferences regarding RA self-efficacy based on a sample of 171 RAs from a large university in the southeastern United States. The proposed 23-item instrument produced reliable results across items and RAs to measure RA self-efficacy. The instrument could be used by other institutions to measure RA self-efficacy, assuming that the job duties align with those of the host institution. However, there is still room for improvement with the existing instrument such as optimizing the measurement of RAs with higher levels of self-efficacy. Insights from this procedure can guide higher education institutions interested in developing instruments for use with student housing staff to inform hiring, training, and intervention programs. Implications for practice and policy are also provided.

. . . .

▌ *Authors*

CALAH K. FORD
Director of the Undergraduate Elementary Education Program and Visiting Assistant Professor Georgetown College
Calah_Ford@ georgetowncollege.edu

CHRISTOPHER WHITE
Assistant Director for Strategic Initiatives in Residence Life Office of the Vice President for Student Success (formerly Student and Academic Life) University of Kentucky
chris.white@uky.edu

ABIGAIL M. A. LOVE
Researcher
Autism Spectrum Australia
love.abigail.a2@gmail.com

KENYATTA JETER
Associate Director of Residence Life Office of the Vice President for Student Success (formerly Student and Academic Life) University of Kentucky
kenyatta.jeter@uky.edu

CAIHONG R. LI
Statistician
Center for Research on Violence Against Women University of Kentucky
caihong.li@uky.edu

NICHOLAS P. KEHRWALD
Dean of Students
University of South Carolina Upstate
Kehrwald@uscupstate.edu

JUSTIN BLEVINS
Director of Residence Life Office of the Vice President for Student Success (formerly Student and Academic Life) University of Kentucky
j.blevins@uky.edu

CORRINE M. WILLIAMS *
Associate Professor in the College of Public Health and Acting Vice President for Student Well-Being in the Office of the Vice President for Student Success University of Kentucky
corrine.williams@uky.edu

TRISHA CLEMENT-MONTGOMERY
Dean of Students
Office of the Vice President for Student Success (formerly Student and Academic Life) University of Kentucky
trisha.clement@uky.edu

> *With growing attention to the influence of the residential experience on student retention, well-being, and satisfaction, it is important for student housing administrators to support and manage both newly hired and returning RAs in this influential position by appropriately measuring RAs' beliefs about their abilities.*

Resident advisors (RAs) serve as leaders, facilitators, and mentors for communities of college students living in campus residence halls within the United States (Boone et al., 2016; Porter & Newman, 2016). Each year, new RAs are hired and trained to support student housing. With growing attention to the influence of the residential experience on student retention, well-being, and satisfaction (Brooks, 2010), it is important for student housing administrators to support and manage both newly hired and returning RAs in this influential position by appropriately measuring RAs' beliefs about their abilities. Little is known about how newly hired and returning RAs perceive their abilities to perform—or, their self-efficacy—in this role, which Taub et al. (2016) have identified to be vital for training.

Perceptions of abilities are most appropriately measured using domain-specific self-efficacy scales to provide more accurate insight into personal beliefs about abilities (Bandura, 2006). Investigations of self-efficacy in various domains in licensed occupations (e.g., Love et al., 2020; Mamaril et al., 2016) suggest that self-efficacy influences performance within these domains. A large body of research has examined occupational self-efficacy, specifically within domains of helping occupations such as teaching and nursing. Zee and Koomen (2016) have found that teachers who reported higher levels of teacher self-efficacy also report higher levels of psychological well-being and improvement in teaching behavior. In nursing, Fida et al. (2018) found that nurses who reported high levels of nursing self-efficacy reported less of the burnout commonly associated with this career. These examples demonstrate the importance of self-efficacy in employees in helping occupations.

Like other professionals in helping occupations, RAs influence the safety and success of others. Researchers have yet to establish the similar role of self-efficacy in the RA position. Utilizing educational measurement practices as applied to helping occupations will inform the well-being and success of RAs. To understand the role that RA self-efficacy plays in daily task performance as well as the hiring and training of RAs, the current study extends occupational self-efficacy research, specifically examining beliefs held by RAs to fulfill their duties.

THE ROLE OF A RESIDENT ADVISOR

RAs are accessible peer leaders in residence halls who serve as university representatives and work to promote successful living and learning environments (Blimling & Schuh, 2015). RAs are typically college-age students who have met qualification criteria (e.g., on-campus residency, grades, conduct) for employment with their institutions' student housing or residence life offices. Not only are these young adults hired as staff, but this opportunity typically serves as a simultaneous leadership development experience for them. Most importantly, RAs are integral in creating a sense of community and belonging for their residents—especially for new students—and are entrusted to act as stewards of the university (Duran et al., 2020; Porter & Newman, 2016).

Typical RA duties include building a sense of belonging, crisis management, and policy administration, and success requires specific interpersonal, communication, and organizational skills (Duran et al., 2020). To assist administrators who hire, train, and evaluate RAs, it is consequential to understand the perceptions of students who are selected to serve in this important position and whether they can steward these spaces with confidence.

MEASURING RA SELF-EFFICACY

Social Cognitive Theory (SCT; Bandura, 1997) posits that human functioning occurs through interactions between one's environmental, individual, and behavioral factors through a process of triadic reciprocation. One of the most studied human functions is self-efficacy, or individuals' belief or confidence in their ability to perform specific tasks under specific circumstances. In the context of this study, human functioning occurs as interactions between environmental (student housing), individual (RA self-efficacy), and behavioral (RA duties) factors. Social Cognitive Theory emphasizes that RAs have the agency to change their perceptions and behaviors while serving in this position.

Self-efficacy influences motivation and effort: aspects of employment that would interest institutions, administrators responsible for hiring and training, and/or staff who support and serve in the RA role. Bandura (2006) recommended that measuring self-efficacy should be specific to the tasks being considered. Although other helping professions, such as counselors or teachers,

seem to share similar responsibilities—such as stewardship, communication, and conflict resolution—with resident advisors, the RA's role is unique because peer students (those of a similar age and developmental stage) are typically hired, and RAs are not required to hold a license or certification to perform their duties. Given the unique roles and expectations of RAs, it is important to carefully consider the specific duties of the position when measuring their self-efficacy.

Empirical investigations regarding the measurement of RA self-efficacy exist, but the only study, to our knowledge, to develop an RA self-efficacy instrument grounded in self-efficacy theory is Denzine and Anderson's (1999). Specifically, they developed items focused on the RA's ability to positively influence student development, administered their instrument to students currently employed as RAs ($N = 111$), and used exploratory factor analysis (EFA). Although EFA based on polychoric correlations (not used by Denzine and Anderson) can be as efficient as Rasch modeling at examining dimensionality (Christensen et al., 2012), EFA does not disclose item and person location on the construct (Schumacker & Linacre, 1996). Moreover, Denzine and Anderson's instrument is not grounded in specific RA job duties.

STUDY PURPOSE AND RESEARCH QUESTIONS

In our study, a locally developed instrument composed of statements sourced from the major job duties established by the host institution was used to measure RA self-efficacy. The data were analyzed to address

> *. . . RAs are integral in creating a sense of community and belonging for their residents—especially for new students—and are entrusted to act as stewards of the university.*

research questions using Rasch measurement (Bond & Fox, 2015) to determine the quality of items that form a meaningful variable. To our knowledge, no other study has applied Rasch analysis to RA self-efficacy. Rasch techniques allow us as researchers to use item-level and instrument-level diagnostic information to optimize the measurement of latent constructs like self-efficacy to differentiate levels of RA self-efficacy. Practitioners (e.g., university and student housing administrators, residence hall directors), on the other hand, can use Rasch findings to inform hiring, training, and evaluation.

Our study was guided by the following research questions (RQs):

1. *Do items that represent RA duties support a unidimensional interpretation of the construct?*

2. *If unidimensional, can the new instrument reliably differentiate distinct groups of RAs?*

3. *Do the response categories support varying degrees of endorsement regarding RAs' perceptions to perform duties?*

METHOD

Sample and Procedure

Participants were from a large university in the southeastern United States. Data were collected in fall 2018 as part of a larger project in partnership with staff in housing and residence life. A total of 215 RAs were administered the measure in this study prior to RA training, about three weeks before fall semester started. The response rate was 80%, with 171 RAs completing the instrument. New RA hires (newly hired, first semester as an RA) made up 43% and female RAs made up 52% of the sample. The racial and ethnic identities of the sample included White or Caucasian at 59%; Black or Afri-

can American at 21%; Spanish, Hispanic, or Latinx at 8%; Asian American or Pacific Islander at 4%; Native American or Alaskan Native at 1%; two or more races at 3%; and undisclosed at 4%.

Resident Advisor Self-Efficacy Instrument

The Resident Advisor Self-Efficacy Instrument (RASE) was developed after a review of the RA job description and self-efficacy literature and was guided by general instrument development principles recommended by Bandalos (2018). The research team originally developed 33 items based on the RA job description. RAs rated their ability to fulfill their expected job duties on a 4-point Likert-type response scale ranging from 1 (*I cannot do this*) to 4 (*I can definitely do this*).

Prior to finalizing the RASE, items were subjected to expert reviews with residence life staff to ensure that they would be interpreted as intended. Administrators from the host institution ($n = 7$) and direct supervisors of RAs ($n = 32$) were sent the instrument and corresponding job description by email. Experts were asked to provide informal (qualitative) feedback on the alignment of items with the original job description and recommendations for improving the item phrasing (e.g., clarifying, collapsing, expanding, or deleting items). This process resulted in 23 items that were subjected to data analysis (see Table 1).

DATA ANALYSIS

The analysis for this study was guided by Rasch measurement theory (Rasch, 1960) to determine the measurement properties and functioning of the RASE. The research team used Winsteps 4.7.1 (Linacre, 2020) to fit data to the Rasch rating scale model (RSM; Andrich, 1978). This model is used for rating scale data where each item format is the same across the instrument. The RSM

TABLE 1

Item Quality Index of RASE Using 4-Point and 3-Point Response Scales

ITEM	STATEMENT	MEASURE	SE	INFIT	OUTFIT
5	Locate appropriate resources in response to difficult situations.	0.92 (0.94)	0.18 (0.18)	1.15 (1.16)	1.19 (1.21)
3	Effectively deal with differences between residents.	0.86 (0.93)	0.18 (0.19)	1.33 (1.23)	1.34 (1.27)
14	Handle crises in an appropriate manner.	0.72 (0.73)	0.19 (0.19)	1.01 (1.02)	1.02 (1.03)
21	Balance personal and professional obligations.	0.69 (0.70)	0.19 (0.19)	1.01 (1.03)	1.05 (1.06)
17	Exercise time management.	0.55 (0.55)	0.19 (0.19)	1.18 (1.19)	1.20 (1.20)
23	Identify ways to improve in my position.	0.47 (0.44)	0.19 (0.19)	1.10 (1.03)	1.03 (1.00)
4	Address violations of the code of conduct appropriately.	0.40 (0.41)	0.19 (0.19)	1.02 (1.03)	0.96 (0.97)
8	Effectively facilitate programming.	0.40 (0.41)	0.19 (0.19)	1.17 (1.19)	1.09 (1.11)
9	Connect residents to community engagement opportunities.	0.26 (0.27)	0.20 (0.20)	1.15 (1.16)	1.19 (1.19)
12	Assist residents in identifying negative behaviors impacting academic success.	0.25 (0.26)	0.19 (0.20)	1.01 (1.01)	0.96 (0.97)
15	Respond to emergencies following department protocol.	0.21 (0.22)	0.20 (0.20)	1.19 (1.20)	1.27 (1.27)
13	Handle incidents in an appropriate manner.	0.10 (0.10)	0.20 (0.20)	0.77 (0.77)	0.64 (0.65)
18	Demonstrate effective organizational skills.	0.06 (0.06)	0.20 (0.20)	0.99 (1.00)	0.94 (0.94)
6	Effectively facilitate one-on-one meetings with residents.	-0.06 (-0.06)	0.20 (0.20)	1.01 (1.02)	0.93 (0.94)
22	Actively engage in professional activities.	-0.23 (-0.23)	0.21 (0.21)	0.74 (0.74)	0.63 (0.63)
7	Effectively facilitate community gatherings.	-0.32 (-0.32)	0.21 (0.21)	0.97 (0.97)	0.83 (0.83)
11	Assist residents in identifying positive behaviors impacting academic success.	-0.54 (-0.54)	0.22 (0.22)	0.79 (0.79)	0.61 (0.61)
16	Complete all administrative responsibilities in a timely manner.	-0.59 (-0.59)	0.22 (0.22)	0.88 (0.88)	0.94 (0.94)
2	Develop trusting relationships amongst the residential community.	-0.69 (-0.69)	0.22 (0.22)	0.80 (0.81)	0.64 (0.65)
20	Maintain open lines of communication with staff members.	-0.74 (-0.74)	0.22 (0.22)	0.77 (0.77)	0.59 (0.59)
19	Maintain open lines of communication with supervisors.	-0.84 (-0.84)	0.23 (0.23)	0.90 (0.90)	0.73 (0.73)
10	Model positive academic behavior for residents.	-0.89 (-0.89)	0.23 (0.23)	0.91 (0.91)	1.40 (1.40)
1	Develop positive relationships with residents.	-1.00 (-1.00)	0.23 (0.23)	0.90 (0.90)	0.76 (0.78)

Note. Values in parentheses represent results for the RASE 3-point response scale.

Infit and outfit based on mean square (MNSQ) fit index.

Table sorted by item quality index measure.

SE = standard error.

allows a researcher to investigate both item and person locations (i.e., RA endorsement) which are placed on the same metric. Specifically, the RSM allows us to understand if the data is unidimensional and determine if items are independent of each other. Addressing RQ1 in this way helps determine whether the items collectively represent a single RA self-efficacy construct and if that single RA self-efficacy construct influences item responses. The RSM helps us investigate RQ2 by identifying how many distinct (reliable) groups of RAs the RASE separates along the RA self-efficacy continuum. Last, the RSM allows us to optimize the measurement of RA self-efficacy by identifying how categories are used within the RASE and how well each item fits the RSM (RQ3).

Unidimensionality and Local Item Independence Assessments (RQ1)

A principal component analysis of the residuals (PCAR) was used to evaluate unidimensionality. Linacre (2020) identified Rasch-based criteria to assess this assumption: Model variance higher than 50%, standardized residuals of the eigenvalue of the first contrast less than 2.0, and proportion of model variance explained was much larger than the unexplained variance in the first contrast. Any concerns regarding unidimensionality were inspected for item clustering at high or low loadings in the first contrast of the standardized residuals. Bond and Fox (2015) reported that this assumption can be held if no concerns were identified through previous analyses of the data. Otherwise, item content analysis was conducted to determine possible construct-relevant differences among items clustering at high or low levels.

Closely related to the unidimensionality assessment, local item independence assessment provides assurance to continue with Rasch analysis. Lord and Novick (1968) established that items should only be correlated by the latent construct (i.e., RA self-efficacy), and no significant correlations should be shared between items after accounting for the latent construct with cutoff value relative to the data, generally ranging from 0.1 to 0.7 (Christensen et al., 2017). The cutoff value established herein is .52 and was based on Marais' (2013) recommendation (i.e., average residual correlation +.2).

Reliability Assessment (RQ2)

Rasch item and person reliability were evaluated, with values above 0.8 expected for person reliability and above 0.9 for item reliability (Linacre, 2020). Reliability, in Rasch analysis, provides information about the reproducibility of the data. Person reliability indicates the degree of separation between respondents at different levels of RA item endorsement. Item reliability indicates item spread, reflecting the RA self-efficacy continuum. The Rasch person separation index, a reliability index that identifies how many statistically distinct levels of RA self-efficacy are distinguished by the items, was calculated.

Model-Data Fit Assessment (RQ3)

The RSM was determined using information weighted (infit) and unweighted (outfit) mean square (MNSQ) values, which indicates how well the data fit the model. Item fit to the RSM was determined if infit and outfit MNSQ values were observed to fall between 0.50 and 1.50 (Peabody et al., 2017). To investigate the item difficulty and person ability locations, a Wright variable map was generated. As an added consideration for model-data fit, a Keyform was also produced to ensure that the proper order of response categories was maintained (i.e., *I cannot do this* to *I can definitely do this*).

Preliminary analysis revealed that there were fewer than 10 responses in the first category, but at least 50 in all others. According to Linacre's (2002) recommendation, at least 10 responses per category are recommended to ensure valid estimates. To maintain the optimal RSM, as well as optimal conditions for practical interpretation of the data, two separate RSM analyses were conducted. The first RSM analysis was based on the original 4-point response categories (4-category) format. The second RSM analysis was conducted on 3-point response categories (3-category), achieved by collapsing the first and second categories (i.e., *I cannot do this* and *I'm not sure that I can do this*). Conceptually and arguably, these two categories may reflect the same level of confidence (i.e., not able to do this). This two-track analysis allows us to determine if the 4-category format provides stable estimates or if a 3-category format is needed to optimize the response format.

RESULTS

Unidimensionality and Local Item Independence Assessments (RQ1)

Results from the PCAR of the 4-category RASE indicate that the Rasch dimension explained 35.60% of the variance (eigenvalue = 12.69), and the unexplained variance in the first contrast was 8.40% (eigenvalue = 3.01). The variance explained was less than the 50% threshold, and the eigenvalue of the first contrast was above the 2.0 threshold, as suggested by Linacre (2020). The results from the PCAR of the 3-category RASE (i.e., collapsing the first and second response categories) indicated that the Rasch dimension explained 35.80% of the variance (eigenvalue = 12.80), and the unexplained variance in the first contrast was 7.60% (eigenvalue = 2.73). The variance explained also meets the thresholds suggested by Linacre (2020).

The first contrast and size of the eigenvalue account for a large percentage of variance explained, followed by a distinctly reduced and smaller secondary contrast and eigenvalue for both the 4-category and 3-category response formats. As recommended by Bond and Fox (2015), a further inspection of the item clustering at the top and bottom of a standardized residual contrast plot was conducted. Inspection of the plot did not show any discernible pattern of high or low loadings. Thus, the unidimensional treatment of the RASE was supported.

Standardized residual correlations of the 4-category RASE ranged from .47 (Items 1 & 2) to -.24 (Items 11 & 20; 14 & 21), which fell below the identified criterion of .52. Standardized residual correlations of the 3-category RASE ranged from .52 (Items 1 & 2) to -.20 (Items 4 & 20; 3 & 10; 7 & 20), which again fell at or below the identified criterion of .52. Therefore, local item independence was supported.

Reliability Assessment (RQ2)

The RASE was able to detect RAs who reported low versus high self-efficacy with person reliability of the 4-category (.84) and 3-category (.84) RASE, as well as item separation reliability of the 4-category (.87) and 3-category (.87) RASE. Person reliability was above the preferred threshold of .80, whereas the spread of the items on this instrument, or the item reliability, approached the preferred threshold of .90 (Linacre, 2020).

Both response category formats produced the same person separation. The person separation of 2.33 was transformed into a strata index (Wright & Masters, 1982) of 3.44 (rounded down to 3.0) with the purpose of determining the number of item strata or index for estimating the spread of persons. In our case, the RASE can estimate three distinct groups of RAs.

Model-Data Fit Assessment (RQ3)

Overall, all items had both infit and outfit MNSQ values between 0.59 and 1.40, which fall within the range determined a priori (see Table 1). Therefore, all 23 items fit the RSM for both 4- and 3-category RASE.

Thresholds. Table 2 shows the threshold information for the 4-category and 3-category RASE. Within the 4-category RASE, the standard errors (SEs) of the thresholds for response categories 2 to 3 (-1.01, $SE = .14$) and 3 to 4 (3.26, $SE = .04$) are smaller than the first response category range due to the increased number of respondents within those latter two response categories. For the 3-category RASE, the standard errors of the thresholds for both response category ranges were more similar ($SE_{1, 2 \text{ to } 3} = .14$; $SE_{3 \text{ to } 4} = .04$).

Wright maps. To determine if statements on the RASE reflected the varying degrees of endorsement by RAs, locations of the items and RA abilities were visually inspected. The Wright maps in Figures 1 and 2 display the item locations along the vertical axis for the 23 items and person locations (i.e., RA endorsement), respectively. As demonstrated in these figures, the distribution of items on the right (indicated by item numbers) did not align with the location of each RA (indicated by "#") and the mean ratings ("M"). In general, items based on the 4- and 3-category RASE tended to target RAs with average to below average self-efficacy based on items falling at or below the mean for RAs. That is, because the mean self-efficacy ability for RAs is higher than the mean of the items, it can be concluded that RAs had an easy time stating that they are confident with the RA job duties as reflected in the items. In general, the Wright maps also show how there are several items covering the same space on the construct (RA self-efficacy), which means that several of the items could be replaced with harder-to-endorse items without increasing the number of items much beyond 23. Moreover, the Wright map in Figure 2 shows that the 3-category RASE items overlap more with the RA locations than do the 4-category RASE items. The larger overlap of items with RA locations indicates that the 3-category RASE measures the sample of RAs better than the 4-category RASE.

TABLE 2

Rating Scale Model Thresholds for RASE Using 4-Point and 3-Point Response Scales

INDEX	4-CATEGORIES				3-CATEGORIES		
	1	2	3	4	1-2	3	4
Observed count	2	52	1,146	2,731	54	1,146	2,731
Observed average measure	1.44	2.07	2.82	4.77	0.91	1.68	3.64
Threshold		-2.25	-1.01	3.26		-2.12	2.12
SE		0.71	0.14	0.04		0.14	0.04

Note. 1 = *I cannot do this;* 2 = *I'm not sure that I can do this;* 3 = *I am pretty sure I can do this;* 4 = *I can definitely do this.*

SE = standard error.

FIGURE 1

Wright Variable Map RASE with 4-Point Response Scale
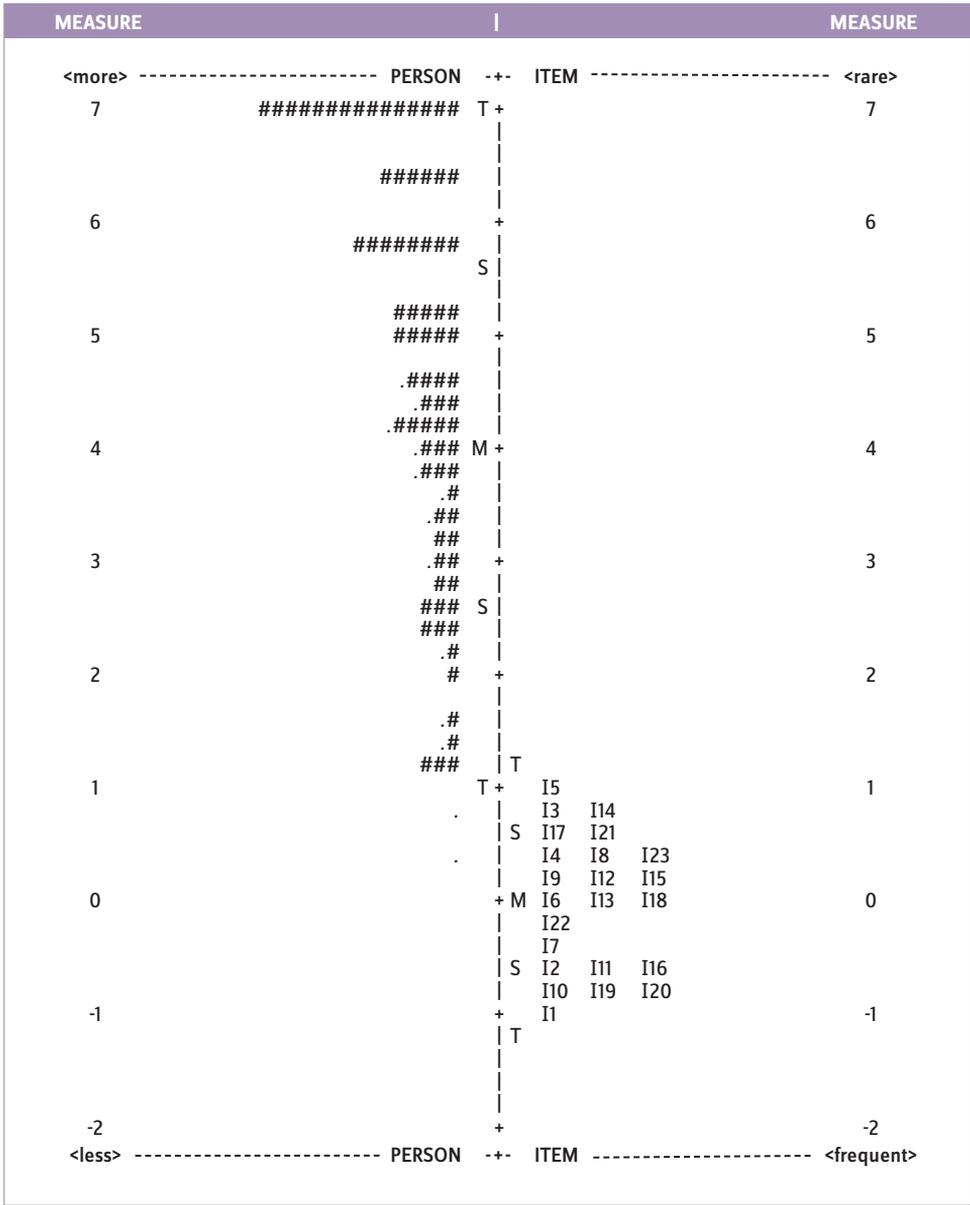
```
MEASURE                            |                        MEASURE

 <more> ----------------------- PERSON  -+-  ITEM ----------------------- <rare>
    7              ###############  T +                                      7
                                      |
                                      |
                        ######        |
                                      |
    6                                 +                                      6
                      ########        |
                                    S |
                                      |
                        #####         |
    5                   #####         +                                      5
                       .####          |
                       .###           |
                       .#####         |
    4                  .### M +                                              4
                       .###           |
                         .#           |
                        .##           |
                        ##            |
    3                   .##           +                                      3
                        ##            |
                       ### S |
                       ###            |
                        .#            |
    2                    #            +                                      2
                                      |
                       .#             |
                       .#             |
                      ###             | T
    1                       T +    I5                                        1
                        .             |    I3    I14
                                    | S  I17   I21
                        .             |    I4    I8    I23
                                      |    I9    I12   I15
    0                                 + M  I6    I13   I18                    0
                                      |    I22
                                      |    I7
                                    | S  I2    I11   I16
                                      |    I10   I19   I20
   -1                                 +    I1                               -1
                                    | T
                                      |
                                      |
                                      |
   -2                                 +                                     -2
 <less> ----------------------- PERSON  -+-  ITEM ----------------------- <frequent>
```

Note. This figure demonstrates the item location of the 23 items of the RASE and the person location for the RAs who responded to the scale. Each "#" in the person column is 2 persons; each "." is 1.

FIGURE 2

Wright Variable Map for RASE with 3-Point Response Scale
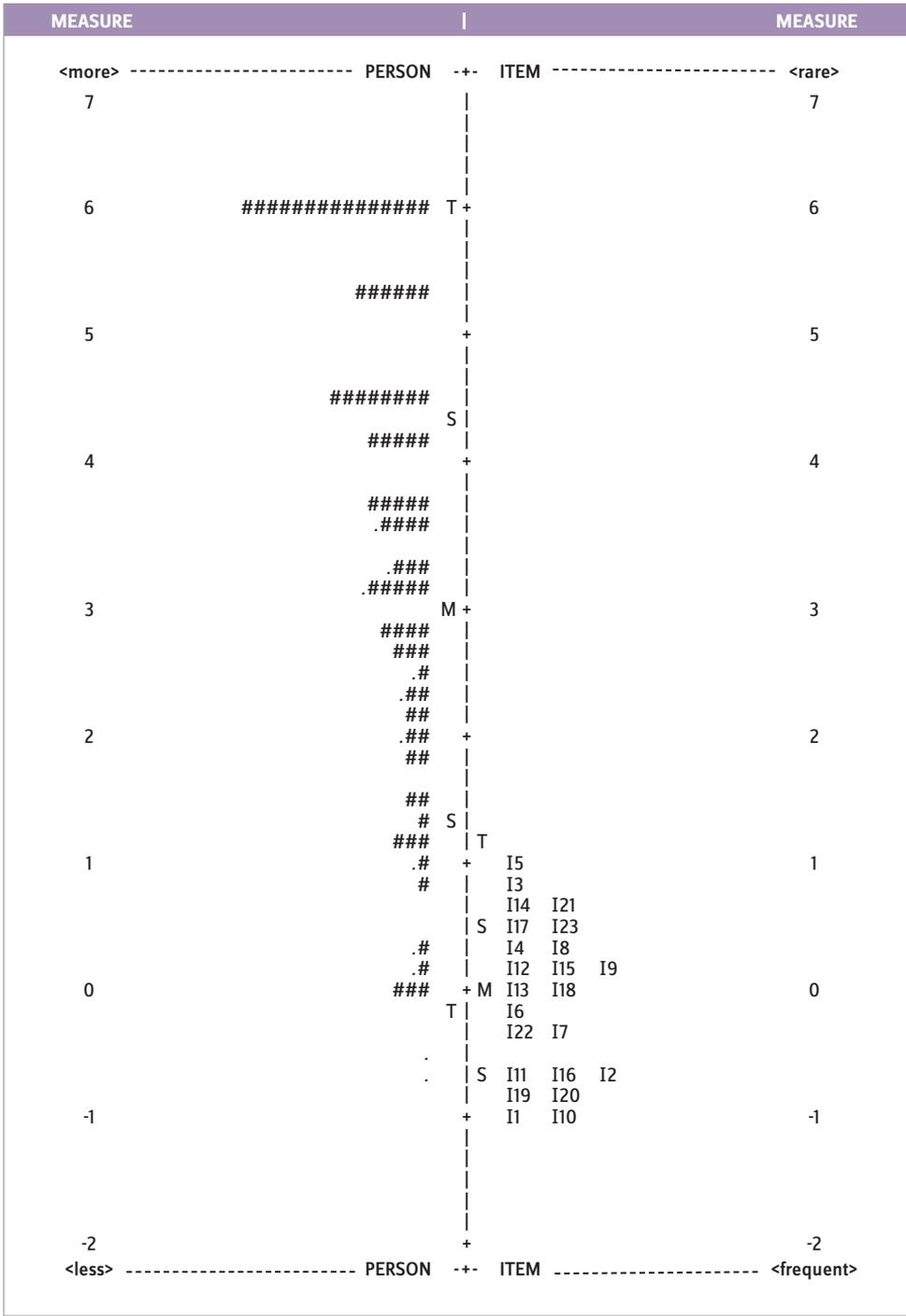
```
  MEASURE                              |                           MEASURE

  <more> ----------------------- PERSON -+- ITEM ---------------------- <rare>
    7                                    |                                7
                                         |
                                         |
                                         |
    6                  ############### T +                              6
                                         |
                                         |
                                         |
                       ######            |
    5                                    +                              5
                                         |
                                         |
                       ########          |
                                   S |
                       #####           |
    4                                    +                              4
                                         |
                       #####            |
                      .####            |
                                         |
                      .###             |
                     .#####            |
    3                               M +                                3
                       ####             |
                       ###              |
                        .#              |
                       .##              |
                        ##              |
    2                  .##               +                              2
                        ##              |
                                         |
                        ##              |
                        #  S |
                       ###     | T
    1                   .#     +    I5                                  1
                         #     |    I3
                               |    I14  I21
                               | S  I17  I23
                        .#     |    I4   I8
                        .#     |    I12  I15  I9
    0                  ###     + M  I13  I18                            0
                              T |    I6
                               |    I22  I7
                               |
                         .       |
                         .     | S  I11  I16  I2
                               |    I19  I20
   -1                          +    I1   I10                           -1
                               |
                               |
                               |
                               |
   -2                          +                                       -2
  <less> ----------------------- PERSON -+- ITEM ---------------------- <frequent>
```

Note. This figure demonstrates the item location of the 23 items of the RASE and the person location for the RAs who responded to the scale. Each "#" in the person column is 2 persons; each "." is 1.

> *One of the most studied human functions is self-efficacy, or individuals' belief or confidence in their ability to perform specific tasks under specific circumstances.*

Keyforms. The visual location of most RAs on the Keyform indicated high levels of self-efficacy (see Figures 3 and 4). The Keyforms also show that item 5 is the most difficult item to endorse, whereas item 1 is the easiest to endorse. Item and response ordering demonstrated expected trends (e.g., Figure 3 response categories were maintained from 1 = *I cannot do this* to 4 = *I can definitely do this*). Most RAs responded to items that are easy to endorse at the highest response category (4). The Keyforms for both 4- and 3-category RASE provided visual and quantifiable evidence that the response categories discriminate between the different levels of endorsements. For the 4-category RASE, Figure 3 visually distinguished three steps (sections marked with ":") and four levels (labeled 1-4) that are appropriately ordered, supporting the use of four response categories for the RASE. Similarly, for the 3-category RASE, Figure 4 presents similar visual representations of the results for this alternative format, but with two steps and three levels that are ordered despite collapsing categories 1 and 2. In addition to the information from the Wright maps, the Keyforms also suggest that the 3-category RASE is a more optimal measure of RA self-efficacy, capturing a closer relationship between item and person locations using a more concise response scale.

There is an issue to note with the visual representation of the data produced through the analyses (Figures 1 and 2). Specifically, items tended to cover the middle range of the RA self-efficacy continuum (-1 to +1 logits which is akin to *z*-scores), but relative to our RA sample most items are easy to endorse. That is, the distribution of RA ability to endorse the items was negatively skewed, which suggested that RAs felt confident in their ability to perform their expected duties, at least as they relate to the job duties examined herein. The implication of these findings is that items on the RASE are most applicable for RAs who do not report a higher ability to endorse RA self-efficacy items. In other words, items may not differentiate RA ability levels for those high on the construct but could be an appropriate instrument for RAs who are low in confidence in their role and need more training.

DISCUSSION

Residential living at higher education institutions is a key component of the typical college or university experience. This experience is especially significant for first-year students. As peer leaders hired to steward residential students, RAs are expected to perform duties identified by their employing institution to the best of their abilities (Duran et al., 2020; Manata et al., 2017). The effectiveness of an RA's performance is directly influenced by their beliefs (Bandura, 1986). The unique role of an RA to meet employer expectations while supporting student residents requires specific and intentional training. This study presented one institution's approach using Rasch measurement to develop an optimal RA self-efficacy instrument for local implementation, providing administrators the opportunity to measure RAs' perceptions about the established job duties.

Validity Evidence for RASE

From this investigation, sufficient evidence was found to support RQ1. Specifically, items on the RASE could be used and scored as representing a single unidimensional

construct. Furthermore, the absence of item misfit and dependence with other items further indicates that the RASE did not present measurement-relevant issues within a Rasch framework. In other words, the measure can be considered useful in measuring RA self-efficacy as an isolated construct. RQ2 was supported given that the RASE can separate out at least three groups of RAs along the unidimensional continuum.

Keyform for RASE with 4-Point Response Scale

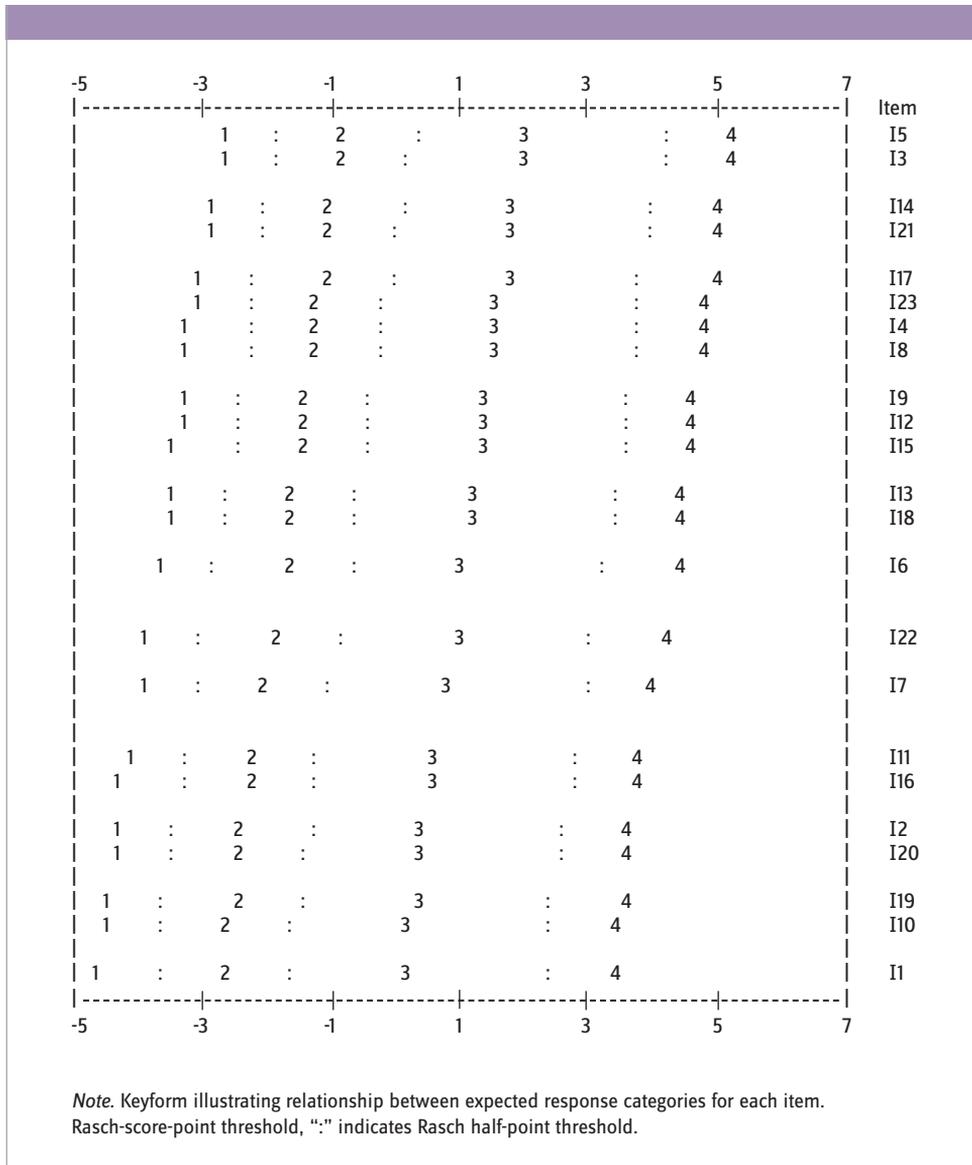*Note.* Keyform illustrating relationship between expected response categories for each item. Rasch-score-point threshold, ":" indicates Rasch half-point threshold.

FIGURE 4

Keyform for RASE 3-Point Response Scale

```
    -5          -3        -1        1         3      5        7
    |-----------+-----------------+-----------------+----------|   Item
    |              1     :            2         :       3        |   I5
    |                                                            |
    |                1     :            2            :     3      |   I3
    |                                                            |
    |                1    :            2           :      3       |   I14
    |                1    :            2          :       3       |   I21
    |                                                            |
    |              1     :            2            :      3       |   I17
    |                                                            |
    |                1   :           2          :        3       |   I23
    |                1   :           2          :        3       |   I4
    |                1   :           2          :        3       |   I8
    |                                                            |
    |               1   :             2        :        3        |   I9
    |               1   :             2        :        3        |   I12
    |               1   :             2        :        3        |   I15
    |                                                            |
    |             1    :            2          :     3           |   I13
    |             1    :            2          :     3           |   I18
    |                                                            |
    |             1   :             2          :     3           |   I6
    |                                                            |
    |           1    :              2            :    3          |   I22
    |                                                            |
    |           1    :              2            :    3          |   I7
    |                                                            |
    |            1    :              2          :      3         |   I11
    |             1    :              2         :      3         |   I16
    |                                                            |
    |           1    :              2           :    3           |   I2
    |           1    :              2           :    3           |   I20
    |                                                            |
    |           1   :              2          :     3            |   I19
    |         1    :              2          :   3               |   I10
    |                                                            |
    |         1    :              2            :    3            |   I1
    |-----------+-----------------+-----------------+----------|
    -5          -3        -1        1         3      5        7
```

Note. Keyform illustrating relationship between expected response categories for each item.
Rasch-score-point threshold, ":" indicates Rasch half-point threshold.

RQ3 was addressed by the two-track analyses of the data using the original 4-point response scale format and 3-point response scale format. Results indicated that both response category scales had similar patterns of model fit, reliability, item quality, and item and person locations. However, the thresholds for the 3-category format RASE provided more optimal results than did the original 4-category scale.

Concerning the Keyforms (Figures 3 and 4) and associated thresholds (Table 2), the person reliability results for both response scale formats indicated that the RASE was able to distinguish between varying levels of RA abilities. Item reliability (.87) approached the preferred threshold (.90), which flagged minor concerns about the spread of the items to reflect the continuum of the construct but ultimately supported the extent of the RASE statements to collectively measure the construct. Reliability results provided sufficient evidence for the appropriate use of the RASE as a local measure of RA self-efficacy, as defined by the major jobs and duties established by the host institution. In addition, reliability results provide evidence that there are quantifiable distinctions (at least three) among RAs who responded to the RASE.

Evidence from these Rasch analyses suggested that the distribution of the person and item means were in different locations. For example, for the analysis of the 4-category RASE, M_{Person} = 4.04; M_{Items} = .00. This misalignment was most obvious upon visual inspection of the Wright maps (see Figures 1 and 2). This extreme response style, also known as a ceiling effect, can be interpreted as construct underrepresentation (Messick, 1996) but was an expected trend, as evidenced by other self-efficacy studies (Chang & Englehard, 2016; Toland & Usher, 2016). One way to overcome this issue in future instruments would be the inclusion of known job duties that would be more difficult for RAs to accomplish (endorse) unless training was provided. Doing so would allow housing administrators to identify potentially problematic job duties and inform future RA training.

Implications for Practice and Policy

Commonly, RAs participate in fall and spring training related to various aspects of their job. The results from the RASE could aid in creating a customized RA training schedule that would select the areas in which the RA needs more support. If an institution facilitates conference-style training, RAs can attend the sessions that best fit their professional development needs. Additionally, RAs can participate in regular one-on-one meetings with their supervisors, where their RASE scores and self-efficacy are discussed. Using the RASE can assist supervisors in having professional development and/or coaching conversations that cater to the needs of the RA. Also, hall directors can utilize the RASE to create professional development opportunities (perhaps during staff meetings) that cover areas of concern that may benefit several of their staff members as indicated in their RASE score.

Although it should not be used as the sole method for selecting candidates, the RASE can be administered during the RA hiring process to gauge the overall self-efficacy of the RA candidate pool and the

. . . the optimal measurement of RA self-efficacy is one that is based on an intentional purpose (i.e., hiring, training, evaluation) and guided by collaborative relationships with experts.

sense of each candidate's self-efficacy relative to the pool. Furthermore, if the RASE is completed by returning and new RAs, hall directors can build a team based on residence hall needs and staff self-efficacy strengths. That is, the RASE can be used to help with placement of RAs within halls or apartments based on job duties they are more and less confident in performing.

Finally, several considerations should be weighed as institutions venture into local measurement or building of their own RA self-efficacy tool. In general, those interested in measuring RA self-efficacy should consider the benefits of an instrument aligned to the job description and use of modern measurement techniques such as the Rasch model. By approaching this task with an instrument developed uniquely for the institution and analyzing data using the Rasch model, researchers and practitioners can mitigate typical issues (e.g., redundant items, sample-related limitations) when these measurement practices are overlooked. Researchers and practitioners can anticipate these issues when measuring RA self-efficacy and determine appropriate strategies to collect reliable data and make valid interpretations.

Limitation and Future Directions

The number of RAs employed during the academic year may be limited to smaller samples than those observed in our study. This is a limiting condition for any type of quantitative study, but, with measurement techniques such as the Rasch rating scale model, it is not a deterring condition. Additionally, statements based on job duties that applicants—especially those that are highly qualified—are expected to perform or for which they demonstrate competency before being hired can be easily endorsed, producing data susceptible to an extreme response style at the high end of the response scale.

The results of our study suggest that the optimal measurement of RA self-efficacy is one that is based on an intentional purpose (i.e., hiring, training, evaluation) and guided by collaborative relationships with experts. In our study, the optimal measurement using the RASE was driven by the practical purpose of measuring the perception of critical job duties coupled with meeting ideal instrument development guidelines. These experts should include both applied partners in student housing and research experts in the field of psychology and psychometrics in order to mitigate any construct and measurement-related issues associated with measuring data on this unique role on university and college campuses. Future studies could consider collecting data from RAs at multiple institutions and adding more challenging-to-endorse items (job duties) to build an instrument with broader impact and scope.

CONCLUSIONS

Practitioners could consider the RASE a good instrument at measuring RA self-efficacy based on how all 23 items fit the RSM and evidence that the RSM assumptions were observed. However, the RASE still does not adequately cover the full range of the RA self-efficacy continuum; in particular, more challenging-to-endorse items are needed to better measure or separate out RAs with higher self-efficacy as observed with our sample of RAs. This phenomenon was also observed by Toland and Usher (2016), who recommended that future self-efficacy "studies should identify more challenging-to-endorse items to better match the range of self-efficacy levels being assessed and to improve the generalizability of findings" (p. 953). Additionally, our results show that a 3-category response scale is reasonable for capturing RA self-reports of their abilities to perform RA-specific job duties and responsibilities.

This study proposed the use of a domain-specific self-efficacy instrument related to an important peer leader position at residential colleges and universities in an attempt to identify the optimal conditions to measure this construct. Utilizing the institution's identified major job duties and responsibilities, collaborating as a team of applied practitioners and education and psychology researchers, and utilizing the Rasch rating scale model is a replicable and straightforward approach to understanding important personnel who may influence student success. For practitioners, the end product may facilitate discussions about how they select, train, and develop RAs to meet key job functions. In conclusion, the research team reinforces the necessity of partnership with content experts and contextual considerations when using statistical information to make decisions.

Findings based on this investigation of the RASE also provide evidence for the scale content, which means that all the items accurately represent our construct of interest. This was likely due to the partnership with the host institution. In our study, administrators in housing and residence life (i.e., hall directors, coordinators) provided expert review feedback to develop the final phrasing of the 23 items. Closely working with the residence life staff gave us the opportunity to develop items specifically focused on the intended purpose and utility of the measure—an intentional decision during the instrument development process that is encouraged for other institutions interested in similarly measuring this construct with their corps of RAs. ■

REFERENCES

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561–573. https://doi.org/10.1007/BF02293814

Bandalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford Press.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Information Age Publishing.

Blimling, G. S., & Schuh, J. H. (2015). *Student learning in college residence halls: What works, what doesn't, and why*. Wiley. https://doi.org/10.1002/9781119210795

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers.

Boone, K. B., Bauman, M., & Davidson, D. L. (2016). The evolution and increasing complexity of the resident assistant role in the United States from colonial to modern times. *The Journal of College and University Student Housing, 42*(3), 38–51. https://www.nxtbook.com/nxtbooks/acuho/journal_vol42no3/index.php#/p/38

Brooks, S. E. (2010). The connection between residence life and first-year student retention at the University of Pittsburgh. *The Journal of College and University Student Housing, 37*(1), 12–22. https://www.nxtbook.com/nxtbooks/acuho/journal_vol37no1/index.php#/p/12

Chang, M., & Engelhard, G. (2016). Examining the teachers' sense of efficacy scale at the item level with Rasch measurement model. *Journal of Psychoeducational Assessment, 34*, 177–191. https://doi.org/10.1177/0734282915593835

Christensen, K.B., Engelhard, G., Jr., & Salzberger, T. (2012). Ask the experts: Rasch vs. factor analysis. *Rasch Measurement Transactions, 26*, 1373–1378. https://www.rasch.org/rmt/rmt263a.htm

Christensen, K. B., Makransky, G., & Horton, M. C. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied Psychological Measurement, 41*, 178–194. https://doi.org/10.1177/0146621616677520

Denzine, G. M., & Anderson, C. M. (1999). I can do it: Resident assistants' sense of self-efficacy. *Journal of College Student Development, 40*, 247–256.

Duran, A., Dahl, L. S., Stipeck, C., & Mayhew, M. J. (2020). A critical quantitative analysis of students' sense of belonging: Perspectives on race, generation status, and collegiate environments. *Journal of College Student Development 61*, 133–153. http://doi.org/10.1353/csd.2020.0014

Fida, R., Laschinger, H. K. S., & Leiter, M. P. (2018). The protective role of self-efficacy against workplace incivility and burnout in nursing: A time-lagged study. *Health Care Management Review, 43*, 21–29. https://doi.org/10.1097/HMR.0000000000000126

Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*, 85–106. https://www.researchgate.net/publication/11372384_Understanding_Rasch_measurement_Optimizing_rating_scale_category_effectiveness

Linacre, J. M. (2020). Winsteps Rasch measurement (Version 4.4.5) [Computer software]. https://www.winsteps.com/winsteps.html

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison-Wesley.

Love, A. M. A., Findley, J., Ruble, L. A., & McGrew, J. (2020). Teacher self-efficacy for teaching students with autism spectrum disorder: Associations with stress, teacher engagement, and student IEP outcomes following compass consultation. *Focus on Autism and Other Developmental Disabilities, 35*, 47–54. https://doi.org/10.1177/1088357619836767

Mamaril, N. A., Usher, E. L., Li, C. R., Economy, D. R., & Kennedy, M. S. (2016). Measuring undergraduate students' engineering self-efficacy: A validation study. *Journal of Engineering Education, 105*, 366–295. https://doi.org/10.1002/jee.20121

Manata, B., DeAngelis, B. N., Paik, J. E., & Miller, V. D. (2017). Measuring critical aspects of the resident assistant role. *Journal of College Student Development, 58*, 618–623. https://doi.org/10.1353/csd.2017.0046

Marais, I. (2013). Local dependence. In K. B. Christensen, S. Kreiner, & M. Mesbah (Eds.), *Rasch models in health* (pp. 111–130). Wiley.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*, 241–256. https://doi.org/10.1177/026553229601300302

Peabody, M. R., O'Neill, T. R., & Peterson, L. E. (2017). Examining the functioning and reliability of the family medicine milestones. *Journal of Graduate Medical Education, 9*, 46–53. https://doi.org/10.4300/JGME-D-16-00172.1

Porter, S. C., & Newman, L. (2016). A brief measure of attitudes toward resident advisors. *College Student Journal, 50*, 107–111. https://www.ingentaconnect.com/contentone/prin/csj/2016/00000050/00000001/art00013

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.

Schumacker, R. E., & Linacre, J. M. (1996). Factor analysis and Rasch analysis. *Rasch Measurement Transactions, 9*(4), 470. http://rasch.org/rmt/rmt94k.htm

Taub, D. J., Johnson, R. B., & Calhoun, D. W. (2016). Introduction to the special issue: Resident assistants and college and university housing. *The Journal of College and University Student Housing, 42*(3), 8–11. https://www.nxtbook.com/nxtbooks/acuho/journal_vol42no3/index.php#/p/8

Toland M. D., & Usher E. L. (2016). Assessing mathematics self-efficacy: How many categories do we really need? *The Journal of Early Adolescence, 36*, 932–960. https://doi.org/10.1177/0272431615588952

Wright, B., & Masters, G. (1982). *Rating scale analysis*. MESA Press.

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research, 86*, 981–1015. https://doi.org/10.3102/0034654315626801

DISCUSSION QUESTIONS

....

1. For those competencies that your campus requires of your RAs, what evidence are you using to assess their ability and growth?

2. How is your campus engaging RA self-efficacy as an integral component of RA competence through training, coaching conversations, professional development opportunities, or other means?

3. This article mentions a specific quantitative tool for measuring RA self-efficacy. Would integrating a similar tool be appropriate on your campus? Alternatively, what are some qualitative methods your campus could use to evaluate RA self-efficacy?

4. Thinking about the competencies your campus requires of your RAs, how might the RAs' social identities impact their self-efficacy in the role?

5. Is there a similar evaluation of self-efficacy that may be relevant for graduate student staff or new professionals on your campus?

6. This article highlighted that even with a good instrument it was difficult to measure higher levels of RA self-efficacy using items tied directly to the baseline competencies. What are some characteristics of RAs with high self-efficacy on your campus? What kind of questions might help you assess their competence and growth?

Discussion questions developed by Ashley Staples, The Ohio State University in Columbus; and Tony Cawthon, Clemson University in South Carolina.