Published in final edited form as:

Eval Health Prof. 2014 March ; 37(1): 71–82. doi:10.1177/0163278713500303.

Factor Structure of the Clinical Research Appraisal Inventory (CRAI)

Britain A. Mills, Raul Caetano, and Arianne E. Rhea

UT Southwestern Clinical and Translational Alliance for Research (UT-STAR), UT Southwestern Medical Center

Abstract

The Clinical Research Appraisal Inventory (CRAI) is a 92 item measure covering 10 domains of research self-efficacy. A known behavioral antecedent, reliable and valid measures of self-efficacy represent a potentially useful tool in the evaluation of research training program efficacy. However, few formal psychometric studies of this instrument exist. Using exploratory factor analysis, we examine the CRAI's dimensional structure in a new sample of clinical research trainees. In contrast to the multidimensional solutions reported previously, CRAI responses in the present sample were unambiguously one-dimensional (as suggested by a dominant single eigenvalue and parallel analysis). This discrepant finding may reflect sample differences in research experience, as unlike previous studies, participants had all already obtained a professional degree. The CRAI's dimensional structure may coalesce into a smaller number of factors as research experience is acquired, and investigators should be mindful of this possibility in future studies of the instrument.

Keywords

Clinical Research Appraisal Inventory; exploratory factor analysis

In 2002, the NIH Roadmap for Medical Research highlighted the need for academic homes to foster clinical and translational research (National Institutes of Health, 2006). In response to this need, numerous research institutions have implemented clinical and translational training programs into their educational curriculum. These programs teach investigators how to translate basic research findings into public health outcomes; however, direct research outcomes (peer-reviewed publications, secured grants, interventions, etc.) can take years to produce. The Clinical Research Appraisal Inventory (CRAI) (Mullikin, Bakken, & Betz, 2007) provides an immediate measure of research self-efficacy, or how well a researcher believes they can perform the activities required in clinical research careers. Self-efficacy theory posits that those who are confident they can complete a task have a better chance of doing so, despite any obstacles that may arise. As a known behavioral antecedent (Bandura, 1977), self-efficacy is a potentially useful construct for research training programs. To the

extent that it is a valid and reliable indicator of future research productivity, it can be used to provide both short and long-term feedback about program effectiveness.

Although the CRAI is used by multiple research institutes participating in the Clinical and Translational Science Award (CTSA) program of the NIH, its length (92 items in the original instrument) may represent an unnecessary burden for program participants. Psychometric investigations of the CRAI's structure (Mullikin et al., 2007; Lipira et al., 2010; Robinson et al., in press) have examined various ways the instrument might be simplified. In this relatively small literature, data have consistently been collected from samples with broad educational backgrounds, including pre-med and pre-doctoral students, post-doctoral trainees, and researchers with MDs or PhDs. Moreover, multiple different reduced item sets have been reported, and complex multidimensional solutions are the norm. For example, Mullikin et al. (2007) reported that an 8 factor structure provided an acceptable account of a reduced 88 item measure, and Lipira et al. (2010) reported a 7 factor structure for a reduced 69 item set. Similarly, Robinson et al. (in press) reported a six factor structure underlying a reduced 12 item measure (derived by taking the two strongest loading items on each factor in an analysis using a larger item set). This degree of multidimensionality complicates simplification efforts, given the known dependence of an instrument's reliability on the number of indicators used to measure it (Nunnally & Bernstein, 1994). However, these are the only published studies of the dimensionality of the CRAI; as a result, there is a need for corroborating psychometric analyses of the CRAI's dimensional structure.

Identifying valid, reliable, and efficient measures of research self-efficacy are also relevant to several recommendations from the recent CTSA National Evaluation Final Report (Frechtling, Raue, Michie, Miyaoka, & Spiegelman, 2012). For example, such measures will facilitate the evaluation of recommended expansions to education and training programs. The CRAI covers areas specifically targeted for expansion, including collaboration and communicating findings to policymakers and the public. As it is already being administered to trainees at multiple institutions with a CTSA, efforts to better understand this instrument's dimensional structure are consistent with recommendations for targeted studies of program components, and it will also help to achieve consensus on how the CRAI should be scored (addressing the need for standardized, shared data collection tools among these institutions). As research self-efficacy is a key correlate of future research productivity, such efforts also indirectly address recommendations to expand the study of outcomes in program evaluation.

In the present manuscript, we report a psychometric analysis of the CRAI's dimensional structure in a new sample of clinical researchers. The objective is to gain a clearer sense of research self-efficacy's dimensionality, in order to inform future efforts to simplify this complex instrument. In particular, we address two limitations of studies to date that may have resulted in unnecessarily complex dimensional representations of research self-efficacy. First, we use item parcels – the average of the individual topics within each of the 10 broad areas – as the unit of measurement in an exploratory factor analysis to infer the structure of the scale as a whole. While our sample size is comparable to Lipira et al. (2010) and smaller than sample sizes in the other two studies, a parceling strategy provides a much more acceptable subject to variable ratio for factor analysis than any psychometric study of

the CRAI, at this writing. Subject to variable ratios in these studies range from 1.4:1 to 3.6:1, which are substantially below most recommended cutoffs. At a given sample size, the likelihood that factors are retained by chance alone increases as the number of variables increases (Nunnally & Bernstein, 1994; see also Velicer & Fava, 1998). Consequently, there is a real possibility that extant studies have overestimated the CRAI's dimensionality, which in turn would lead researchers to retain potentially redundant and unnecessary items in a measure that already imposes a substantial burden on respondents. Moreover, by aggregating only across items *within* a domain, a parceling strategy increases the reliability of the individual units of measurement and avoids problems that can arise when parceling items from different dimensions (Little et al., 2002).

Second, studies of the CRAI reviewed above relied to varying extents on the Kaiser-Guttman "eigenvalue greater than one" rule to determine the number of factors to extract. Given the tendency of this rule to overestimate construct dimensionality (Nunnally & Bernstein, 1994), we instead use the more contemporary method of parallel analysis (Horn, 1965; Humphreys & Ilgen, 1969; Humphreys & Montanelli, 1975) to determine the number of factors. This approach involves generating a series of matrices using randomly generated data based upon the same number of "subjects" and "variables" as in the real data matrix. Each random matrix is then factored, and the eigenvalues derived from each individual analysis are averaged, yielding estimated eigenvalues that would be expected to be obtained by chance, assuming no underlying structure. If a real eigenvalue for a factor exceeds the simulated eigenvalue, that factor represents something beyond random fluctuations in the data. Compared to the Kaiser-Guttman rule, parallel analysis is known to produce more parsimonious – and accurate – estimates of construct dimensionality.

Method

Participants

This study sample included 120 scholars and trainees who participated in the UT Southwestern Master of Science in Clinical Science graduate program between 2007 and 2012. The graduate program provides two educational tracks. The clinical scholars program leads to a Masters of Science in Clinical Science with Distinction, and requires submission of a grant application, in addition to the curriculum and research practicum requirements of the trainee program that leads to a Masters of Science in Clinical Science. Education program participants currently encompass the entire study population for the CRAI. Completion of the CRAI is voluntary and is part of a larger battery of assessments completed annually. Students are advised of evaluation activities during program orientation; they are also informed that evaluation results will facilitate program improvements and outcomes research, benefiting all institutions with a CTSA. Invitations to complete the CRAI are sent electronically each year in July.

Between 2007 and 2012, 120 participants were given at least one opportunity to complete a CRAI assessment. Of these participants, 68 completed at least one survey, giving an overall response rate of 57%. This sample consisted of 29 men and 39 women. Participants had a mean age of 37 years old and were an average of 9 years removed from their terminal degree. The sample was 68% Caucasian, with 84% of the sample having Non-Hispanic

ethnicity. All participants held a terminal (professional) degree: 57 had MDs without a PhD, 4 had PhDs without an MD, 3 had MD/PhDs, 2 had PharmDs, 1 had a DDS, and 1 had a DO. Other common, secondary degrees included the MPH (3), RN (2) and MS (2). Corresponding to the mean age of 37, 76% of the sample represented early-career investigators, holding instructor, fellow, or assistant professor positions.

Measures

The 92 item CRAI is scored on an 11-point scale, where 0 signifies no confidence and 10 signifies total confidence in a skill relevant to clinical research. The instrument covers 10 domains: conceptualizing a study (10 items); designing a study (12 items); collaborating with others (8 items); funding a study (10 items); planning and managing a research study (11 items); protecting research subjects and responsible conduct of research (11 items); collecting, recording, and analyzing data (11 items), interpreting data (4 items); reporting a study (12 items); and presenting a study (3 items). Parcels were created by taking the mean of items within each domain, providing 10 total items used in subsequent factor analyses.

Procedure

The CRAI is administered annually to each program participant, regardless of educational track or year of entry. Dissemination occurs electronically through REDCap (Research Electronic Data Capture) software (Harris et al., 2009). Every student has the opportunity to complete the CRAI at least 3 times during their enrollment (unless the person withdraws or takes a leave of absence from the program). Because this study focused strictly on the CRAI's psychometric structure, analyses were restricted to participants' first observed completion of the CRAI. Exploratory factor analytic models were fit using Mplus 7 (Muthén & Muthén, 2010) using a maximum likelihood estimator. Given previous reports of multifactor structures underlying the CRAI, we examined both one-dimensional and multidimensional solutions. Following the initial extraction, we first used parallel analysis to examine dimensionality by comparing extracted eigenvalues with those that would be expected by chance alone (i.e., assuming no underlying structure; Fabrigar et al., 1999). Solutions were also compared in terms of their theoretical sensibility. For multidimensional solutions, oblique rotations were always examined first to ensure no inappropriate orthogonality constraints were imposed on factor correlations. These solutions were rotated using Mplus' oblique rotation algorithm GEOMIN. All participants provided informed consent and the study was approved by the University of Texas Southwestern Medical Center's IRB.

Results

Bivariate correlations between the 10 parcels are shown in Table 1. As shown in the scree plot in Figure 1, the extracted eigenvalues from the unrotated solution were characterized by a large, dominant single eigenvalue, indicative of a one dimensional structure. Consistent with this finding, eigenvalues for subsequently extracted factors consistently fell below those generated by random data in the parallel analysis, suggesting these factors may simply reflect chance fluctuations in the data. Loadings from this one dimensional structure can be found in Table 2. Magnitudes of the loadings were generally comparable and ranged from .

71 to .91. with “Collecting, recording, and analyzing data,” “Designing a study,” and “Conceptualizing a study” representing the three items with the strongest loadings.

We also examined solutions with up to six factors and present results from these analyses in the second half of Table 2. Extracting two and five factors resulted in models that failed to converge, and multiple loadings were nonsignificant in the four and six factor solutions. Consequently, we report findings for the three factor oblique solution. As seen in Table 2, all but one item – “Planning and managing a research study” – loaded on the first factor. Three items loaded significantly on the second factor, but of these, “Planning and managing a research study” had a substantially higher loading than “Funding a study” or “Collecting, recording, and analyzing data,” both of which also had substantial cross-loadings on the first factor. The third factor reflected a “singlet” dimension comprised of one item, “Presenting a study,” which also cross-loaded heavily on the first factor. Structural estimates for the three dimensional oblique solution also indirectly supported the one factor model, as all items correlated strongly with the first dimension.

Discussion

The original 92 items of the Clinical Research Appraisal Inventory reflect 10 *a priori* subdomains of research self-efficacy. Although extant psychometric studies suggest the CRAI’s underlying dimensional structure may be simpler (ranging from six to eight dimensions), these structures still reflect a complex multidimensional construct. In the present study, we found evidence of a much simpler, one-dimensional structure. Apart from the possibility that factors were simply over-extracted due to the use of liberal extraction techniques, what might explain this discrepancy?

One notable difference between our study and previous findings concerns our use of an item parceling strategy. In a thorough review of literature on the pros and cons of parceling, Little et al. (2002) noted that the dimensionality of a construct can seriously affect the accuracy and validity of parceling techniques. In particular, creating a parcel from a set of items that reflects multiple underlying constructs can potentially obscure invalid dimensional assumptions, particularly when it is unintentional (there are however sensible, intentional uses of such “domain representative” parcels; see Kishton and Widaman, 1994). However, note that in the present study, we parceled items *within* the 10 originally specified subdomains, rather than across domains. Consequently, the resulting parcels retained their originally theorized domain-specificity.

A reasonable counter to this argument is that the originally theorized 10-factor structure is an inappropriate reference model, as subsequent psychometric studies have shown that the CRAI items are better represented by a 6 to 8 factor solution (Lipira et al., 2010; Mullikin et al., 2007; Robinson et al., in press). In particular, creating parcels representing the original 10 *a priori* domains may obscure underlying multidimensionality, if the true factors reflect a mixture of items across these domains. However, the findings reported in those studies suggest this was not the case. Specifically, the simplified factors reported to date entirely reflect *collapsed* domains from the original 10-domain structure – rather than mixtures of items from different domains. As a result, it is unlikely that a parceling strategy obscured

any multidimensionality. To take a specific example, note that items for interpreting, reporting, and presenting results reflected three distinct domains in the original 92 item instrument, but loaded on a single factor in Lipira et al. (2010). If the Lipira et al. (2010) representation of “interpreting/reporting/planning” as a single factor is correct, then our analysis effectively uses three indicators of this construct: one interpreting parcel, one reporting parcel, and one planning parcel. Critically, no assumptions in our modeling approach would have prevented these items from clustering together on a distinct factor.

A more plausible explanation of our finding of a simpler, one dimensional structure is that our sample differed substantively from samples used in other analyses. Comparing characteristics of our sample with those in Lipira et al. (2010), Mullikin et al. (2007), and Robinson et al. (in press), our sample is distinct in that all respondents had already obtained some type of professional research degree. This has several theoretical and practical implications. Theoretically, it suggests that the dimensionality of research self-efficacy may change as a researcher acquires more training and experience. With this experience, fine-grained distinctions in confidence for performing diverse research activities may cease to be drawn, and confidence instead may become increasingly dominated by a global sense of confidence in research skill.

Practically, temporal shifts in dimensionality complicate psychometric assessment. Fortunately, this can be accommodated at the stage of data analysis by techniques such as longitudinal factor analysis. However, it also complicates instrument administration, as the answer to the question of whether a simplified instrument can be delivered may depend on the population in question. In one extreme scenario, complex instruments may be necessary for clinical researchers in early stages of training, while simplified item sets may be more appropriate only in later stages of training. Of course, an instrument’s measurement properties reflect only part of the story, and additional validation analyses relating CRAI subdomains to important criterion variables are needed. For example, early in training, it is unlikely that each of these domains would be equivalent in their ability to predict future research productivity. Validation analyses relating subdomains to indices of productivity (e.g., secured grants, peer-reviewed research publications) would be particularly valuable in determining whether any CRAI item (or subdomain) could be dropped without sacrificing predictive validity.

In sum, research self efficacy represents an intermediate index of future research productivity that can be used in the ongoing monitoring and evaluation of research training programs. The current findings expand the body of knowledge regarding the research and practice of evaluation in clinical and translational science by showing that the CRAI - a common measure of research self-efficacy used in the evaluation of clinical research training programs – has a remarkably simple dimensional structure among clinical research trainees with an advanced degree. To improve the efficiency of clinical research training program evaluation, analyses of item-level (rather than parcel-level) responses to the CRAI using larger, similar samples are therefore recommended, as it is likely that this 92-item instrument can be substantially simplified during later stages of research training.

Acknowledgments

This research study was supported by the National Institutes of Health (NIH) Grant UL1TR0000451 to the University of Texas Southwestern Medical Center.

References

- Bandura A. Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*. 1977; 84:191–215. [PubMed: 847061]
- Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*. 1999; 4:272–299.
- Frechtling, J.; Raue, K.; Michie, J.; Miyaoka, A.; Spiegelman, M. The CTSA National Evaluation Final Report. Rockville MD: Westat; 2012. p. 25-37. Retrieved from https://www.ctsacentral.org/sites/default/files/files/CTSANationalEval_FinalReport_20120416.pdf
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*. 2009; 42:377–381. [PubMed: 18929686]
- Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965; 30:179–185. [PubMed: 14306381]
- Humphreys LG, Ilgen DR. Note on a criterion for the number of common factors. *Educational and Psychological Measurement*. 1969; 29:571–578.
- Humphreys LG, Montanelli RG. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*. 1975; 10:193–206.
- Kishton JM, Widaman KF. Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement*. 1994; 54:757–765.
- Lipira L, Jeffe DB, Krauss M, Garbutt J, Piccirillo J, Evanoff B, Fraser V. Evaluation of Clinical Research Training Programs Using the Clinical Research Appraisal Inventory. *Clinical and Translational Science*. 2010; 3:243–248. [PubMed: 21442017]
- Little TD, Cunningham WA, Shahar G, Widaman KF. To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*. 2002; 9:151–173.
- Mullikin EA, Bakken LL, Betz NE. Assessing Research Self-Efficacy in Physician-Scientists: The Clinical Research Appraisal Inventory. *Journal of Career Assessment*. 2007; 15:367–387.
- Muthén, LK.; Muthén, BO. Mplus (Version 7). Los Angeles: Muthén & Muthén; 2012.
- National Institutes of Health. NIH Roadmap for Medical Research: Fact Sheet. 2006 Retrieved from opasi.nih.gov/documents/NIHRoadmap_FactSheet_Aug06.pdf.
- Nunnally, JC.; Bernstein, IH. *Psychometric theory*. 3. New York: McGraw-Hill; 1994.
- Robinson GFWB, Switzer GE, Cohen ED, Primack BA, Kapoor WN, Seltzer DL, Rubio DM. A Shortened Version of the Clinical Research Appraisal Inventory: CRAI-12. *Academic Medicine*. in press.
- Velicer WF, Fava JL. Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*. 1998; 3:231–251.

Figure 1. Extracted and parallel analysis eigenvalues from an exploratory factor analysis of CRAI item parcels.

Correlations Among CRAI Parcels

Table 1

Domains	1	2	3	4	5	6	7	8	9	10
1 Conceptualizing a study	1.00									
2 Designing a study	.81	1.00								
3 Collaborating with others	.79	.78	1.00							
4 Funding a study	.73	.73	.74	1.00						
5 Planning and managing your research study	.69	.68	.74	.79	1.00					
6 Protecting research subjects/responsible conduct	.59	.63	.59	.56	.64	1.00				
7 Collecting, recording, and analyzing data	.74	.86	.75	.72	.82	.69	1.00			
8 Interpreting data	.68	.76	.62	.54	.48	.59	.71	1.00		
9 Reporting a study	.78	.74	.75	.68	.67	.57	.78	.72	1.00	
10 Presenting a study	.74	.51	.60	.62	.55	.57	.59	.57	.74	1.00

Table 2

Loadings and structural estimates from an exploratory factor analysis of 10 CRAI domain parcels

Domain	3 Factor Solution with an Oblique Rotation									
	1 Factor Solution					Factor Structure				
	1	2	3	1	2	3	1	2	3	
Conceptualizing a study	.88	.82	.07	.14	.89	.60	.31			
Designing a study	.90	1.01	-.02	-.28	.95	.56	-.09			
Collaborating with others	.87	.76	.13	.00	.84	.60	.17			
Funding a study	.83	.62	.27	.06	.80	.66	.22			
Planning and managing a research study	.82	.01	1.19	-.01	.74	1.19	.17			
Protecting research subjects/responsible conduct	.71	.58	.16	.09	.69	.53	.22			
Collecting, recording and analyzing data	.91	.75	.25	-.10	.89	.70	.08			
Interpreting data	.77	.89	-.14	-.02	.80	.41	.13			
Reporting a study	.86	.79	.04	.19	.85	.56	.35			
Presenting a study	.71	.61	-.01	.66	.73	.47	.78			

Note. Bolded loadings are significant at $p < .05$. In the oblique solution, factors 1 and 2 correlated .62, 1 and 3 correlated .19, and 2 and 3 correlated .15. The structural estimate of 1.19 for "Planning and managing a research study" (which should not exceed 1.0) reflected one of two anomalous local parameter estimates for the 3 factor model; the other was a negative residual variance estimate for the same item.