Routledge
Taylor & Francis Group

Journal of Couple & Relationship Therapy
Innovations in Clinical and Educational Interventions

ISSN: 1533-2691 (Print) 1533-2683 (Online) Journal homepage: https://www.tandfonline.com/loi/wcrt20

Testing the Gay and Lesbian Relationship Satisfaction Scale with Item Response Modeling

Christopher K. Belous, Richard S. Wampler & Britney L. Ledford

To cite this article: Christopher K. Belous, Richard S. Wampler & Britney L. Ledford (2020): Testing the Gay and Lesbian Relationship Satisfaction Scale with Item Response Modeling, Journal of Couple & Relationship Therapy, DOI: 10.1080/15332691.2020.1746460

To link to this article: https://doi.org/10.1080/15332691.2020.1746460

Published online: 15 Apr 2020.

JOURNAL OF COUPLE & RELATIONSHIP THERAPY
https://doi.org/10.1080/15332691.2020.1746460

Testing the Gay and Lesbian Relationship Satisfaction Scale with Item Response Modeling

Christopher K. Belous[a], Richard S. Wampler[b,c], and Britney L. Ledford[d]

[a]Purdue University Northwest; [b]Michigan State University; [c]Texas Tech University; [d]Peachford Hospital

ABSTRACT

The Gay and Lesbian Relationship Satisfaction Scale (GLRSS) was developed to measure relationship satisfaction and social support in same-gender couple relationships. Originally validated with traditional Classical Test Theory psychometrics, this study examined the items themselves more closely by subjecting the data to an Item Response Modeling analysis—specifically utilizing Samejima's Unidimensional Graded Response Model. Results indicate that each of the subscales of the measure (Relationship Satisfaction and Social Support) do in fact measure the traits intended, providing evidence of content and construct validity. This study further validates the GLRSS as a valuable and reliable tool for use by researchers and clinicians wishing to study same-gender couple relationships.

KEYWORDS

Item response modeling; same gender couples; assessment; relationship satisfaction

Compared to the number of scales used to determine relationship satisfaction for heterosexual couples, scales and instruments used to identify satisfaction in gay and lesbian relationships are very limited (Belous & Wampler, 2016). Of the U.S. adult population, 3.5% self-identify as gay, lesbian or bisexual (Gates, 2011). Looking more closely at this population, recent estimates indicate up to 46% of gay men and 62% of lesbian women (ages 18–59) in same-gender relationships are cohabitating (Carpenter & Gates, 2008). Previously, relationship scales developed for heterosexual couples such as the Dyadic Adjustment Scale (DAS; Spanier, 1976) and the Revised Dyadic Adjustment Scale (RDAS; Busby et al., 1995) have been used to measure relationship satisfaction in same-gender couples. However, the assumption that instruments designed for heterosexual couples can be used appropriately to assess those in same-gender relationships can be unintentionally discriminatory. The belief that same gender couples face the same levels of stress or public stigma associated with their relationships as different gendered couples is inherently minimizing of their lived experiences, similar to the way in which any measure of assessment should be

CONTACT Christopher K. Belous ckb@pnw.edu 2200 169th St, IBCC-CFTC, Hammond, IN 46323.

considered critically before getting applied to a specific clinical population. The Gay and Lesbian Relationship Satisfaction Scale (GLRSS) fills a need to better assess relationship satisfaction with same-gender couples (Belous & Wampler, 2016). Often, same-gender couples are affected by daily struggles and influences on their individual and couple relationship connected with stigma, secrecy, and lack of family support and social acceptance associated with their identities (Belous, 2015; Belous & Wampler, 2016; Kuyper & Fokkema, 2011; Meyer, 2003).

In addition to external stress factors, individuals that identify as lesbian or gay may also deal with internalized homophobia and homonegativity which may lead to a negative sense of self (Malyon, 1982). The resulting conflict between their internalized values and beliefs and their sexual expression desires and intimacy needs can impact their intimate relationships as well as their familial and social interactions throughout their life (Frost & Meyer, 2009). Because of the increase in stress both internally and externally, it may come as no surprise that lesbian and gay individuals account for a disproportionately larger percentage of clients in outpatient mental health treatment compared to the total population (Belous, 2015; Belous et al., 2015; Belous & Wampler, 2016). Through life stresses both internal and external as well as stress specifically related to gay and lesbian individuals, there is a need for an instrument that was designed with sensitivity to these issues. The GLRSS is able to assess relationship satisfaction successfully with items designed specifically for same-gender couples, which in turn provides more accurate information for researchers and providers hoping to successfully address issues in individual or couple therapy.

The initial development of the gay and lesbian relationship satisfaction scale

The authors investigated scales for use with same-gender couples, and only one had been developed specifically for use with this population, the Relationship Assessment Measure for Same Sex Couples (RAM-SSC; Burgoyne, 2001). This measure was a re-write of the Waring Intimacy Questionnaire (WIQ, Waring & Reddon, 1983; Waring, 1984), which was developed and standardized with heterosexual married couples. While Burgoyne did an excellent job of adjusting the wording of the items, it was still completely based on heterosexual norms and standards for relationship satisfaction—and included 90 true-false statements. In order to more appropriately represent the relationship dimension for same-gender couples, these original 90 items were sorted by a sample of younger adults into relationship categories, then subjected to content analysis of topic areas related to gay and lesbian relationship studies (Belous & Wampler, 2016).

A final group of 30 items were developed around the topics of social support and relationship satisfaction—upon initial testing and classical test theory psychometric testing, this item pool shrank to a final total of 24 items that held strong internal consistently and showed evidence of stable validity (Belous & Wampler).

The analysis for the original development of the GLRSS included 295 total participants recruited online through a variety of means (Belous & Wampler, 2016). The vast majority were non-Latinx white persons. Slightly over half identified as lesbian-female (50.5%). Of the 295, 20 were removed because they did not meet criteria, resulting in a total sample size of 275 respondents, sufficient for the type of factor analysis being performed (Cook & Hatala, 2015). Respondents lived across the United States; however, many came from the Midwest area (approximately 45%). The average length of time since "coming out" was 11.5 years. A more detailed breakdown of the sample characteristics can be found in the original article, *Development of the Gay & Lesbian Relationship Satisfaction Scale* (Belous & Wampler, 2016).

In the earlier study, the GLRSS data were subjected to an Exploratory Factor Analysis, followed by a Confirmatory Factor Analysis once a stable structure was identified. The final version of the GLRSS included two stable factors, labeled Relationship Satisfaction (16 items) and Social Support (8 items). The two factors accounted for approximately 35% of the variance, and had eigenvalues over 2.0, with no other factors having a stable structure or cross loadings with items. Based on the loadings in the factor analysis, items loading primarily on the Relationship Satisfaction factor were summed to create a subscale score; items loading primarily on the Social Support factor were summed for a second subscale score. The two subscales were summed to obtain a total GLRSS score. The GLRSS can be scored by summing totals for each subscale. The Relationship Satisfaction subscale includes items 1–16; the Social Support subscale includes items 17–24. Items 1, 2, 9, 14, 16, and 17 are reverse scored (see Table 1 for items in the scale).

Study aims

The purpose of this study is to further validate the Gay and Lesbian Relationship Satisfaction Scale with a more stringent and robust method of providing evidence for construct and content validity utilizing Item Response Theory to conduct an Item Response Model Analysis. While the original scale was shown to have strong evidence that it was measuring and gathering data related to relationship satisfaction and included information on social support in a reliable and valid way, that initial report is not able

to go into detail as to how much of these traits were being measured due to the inherent inability of Classical Test Theory to determine trait level data. This follow up secondary analysis of the data will be able to provide more evidence of construct and content validity, and determine the amount of information/trait data that is possible with the scale.

Methods

Participants & data collection

In this paper we are reporting on secondary analysis from data collected during the creation of the Gay and Lesbian Relationship Satisfaction Scale (Belous & Wampler, 2016). The GLRSS was validated and tested for reliability utilizing a Classical Test Theory (CTT) process, using factor analysis, reliability testing, and gathering evidence of validity in comparison to other measures (criterion validity). This study will further develop construct and content validity of the measure by subjecting the data to more intense and scrutinous analysis through Item Response Theory and Modeling.

Item response theory and modeling

Item Response Theory (IRT)

Item response theory approaches are considered superior to CTT approaches because IRT allows greater flexibility of item construction to allow for greater variance, as well as providing more information from each individual item than is possible with CTT (Hambleton & Jones, 2005). Because IRT requires more complex statistical analyses, it did not become popular until the widespread availability of computers in the 1980s and the development of more sophisticated statistical techniques and packages (Hambleton & Jones, 2005). IRT focuses on each item in the test; thus, it gives more detailed information on each item. It allows for the use of Likert-scale type questions, as well as agreement-based (agree vs. disagree) responses to items (Hambleton & Jones, 2005). This allows for greater response variability within an instrument and for a richness of detail that is often lacking in CTT-assessed items (Baker, 2001). The main focus of IRT models of analysis is on the probability of answering a question in the "correct" way, i.e., to what degree does the item reflect an underlying trait?

Procedures of Statistical Analysis with Item Response Modeling (IRM)

IRM is the statistical approach to IRT. IRM analysis provides evidence of item-level fit characteristics that reflect a presumed latent trait, referred to

as theta (θ, Embretson & Reise, 2000). There are several models available for analysis within IRM; however, the model that best fit the data collected for this study is the Graded Response Model (Allen & Yen, 2002; Samejima, 1969, 2016). The Graded Response Model is capable of estimating probabilities based on Likert-type items that are "graded" or categorized into related outcomes (e.g., "strongly disagree" to "strongly agree"). IRM does not have cut-and-dried rules of interpretation. There are several schools of thought as to how to use data provided through the model. However, most psychometricians agree that IRM should be used as a method of understanding how well an item is "behaving" or "predicting" the latent trait (θ) that the scale is attempting to measure. IRM can also be used as a data reduction technique, as it is mathematically similar to factor analysis.

Specifically, a unidimensional Graded Response Model was used in the analysis of each of the two subscales of the GLRSS. When using an Item Response Model to analyze an instrument, there are five steps to take: (1) examining slope discriminants, (2) comparing category thresholds, (3) noting significant outcomes on the chi-square test, (4) inspecting the graphical output of data (item information curves, test information curves, etc.), and (5) comparing and contrasting all results. The software program IRTPRO 2.1 was used for calculation and estimation of the model (Cai et al., 2011). IRTPRO 2.1 provides item-level characteristics, as well as providing graphical representation of the information in each item and, in a single graph, for all the items tested (in this case, a subscale).

Results

Relationship satisfaction

Examining slope discriminants

Table 1 lists the slope discriminants (A) for all items on the Relationship Satisfaction subscale, as well as the category thresholds (c) and chi-square diagnostics (x^2). A slope discriminant value higher than 2 is considered best, 1 is good, and above .75 is acceptable (Nguyen et al., 2014; Yang & Kao, 2014). The initial examination of the slope discriminant values calculated for the Relationship Satisfaction subscale indicated that the items were able to distinguish the latent trait (θ) being measured (relationship satisfaction) at an acceptable level ($A > .75$). Of the 16 items in Relationship Satisfaction, 10 have a slope discriminant above 1.0, indicating a higher overall ability to distinguish the latent trait. The 6 items below 1.0 are still acceptable, with the lowest slope discriminant being Item 10, "*During our arguments, I never put down my partner's point of view*" ($A = .76$).

Table 1. Item level parameter estimates and diagnostic statistics–GLRSS relationship satisfaction.

Item	Item text	a	c_1	c_2	c_3	c_4	c_5	c_6	x^2	Probability
1	There are some things about my partner that I do not like	1.08	2.88	1.03	−1.09	−1.61	−2.16	−3.52	108.74	.2151
2	I wish my partner enjoyed more of the activities that I enjoy	1.10	3.69	1.98	.70	−.09	−.73	−2.49	152.34	.0156
3	My mate has the qualities I want in a partner	2.02	6.97	5.95	4.51	3.23	.144	–	60.39	.0409
4	My partner and I share the same values and goals in life	1.60	5.87	4.34	3.67	3.33	1.87	−.44	55.51	.7074
5	My partner and I have an active social life	.80	4.58	3.11	1.89	1.26	.15	−1.37	134.87	.0067*
6	My partner's sociability adds a positive aspect to our relationship	1.29	4.94	3.07	2.39	1.63	.76	−1.21	102.40	.0631
7	If there is one thing that my partner and I are good at, it's talking about our feelings with each other	1.72	5.23	3.80	2.73	2.04	.48	−1.04	85.39	.3472
8	Our differences of opinion lead to shouting matches	1.19	4.26	3.02	1.89	1.25	.51	−.89	105.18	.2228
9	I would lie to my partner if I thought it would "keep the peace"	.80	4.17	2.51	1.13	.78	.18	−1.30	118.87	.1849
10	During our arguments, I never put down my partner's point of view	.76	4.00	2.54	1.14	.55	−.29	−1.96	134.14	.0586
11	When there is a difference of opinion, we try to talk it out rather than fight	1.69	7.07	5.89	4.00	2.98	1.52	−1.21	92.48	.0174
12	We always do something to mark a special day in our relationship, like an anniversary	.98	4.52	3.33	2.55	2.05	1.23	−.21	94.94	.0930
13	I often tell my partner that I love him/her	1.54	6.86	5.19	4.66	3.90	2.87	1.50	46.40	.3334
14	Sometimes sex with my partner seems more like work than play to me	1.16	4.92	2.79	1.82	1.16	.39	−.59	119.22	.0921
15	I always seem to be in the mood for sex when my partner is	.78	3.33	2.22	1.09	.44	−.29	−1.74	161.31	.0099*
16	My partner sometimes turns away from my sexual advances	.89	2.99	2.06	.67	.22	−.21	−.160	128.99	.1041

*x^2 significant at $p < .01$.

Comparing category thresholds

Category thresholds, denoted in Table 1 as c_n , indicate the item's ability to distinguish a respondents' likelihood of choosing one scale category over the adjacent category (e.g., *1* vs *2*, or *somewhat agree* vs *agree*.). In the Graded Response Model, c_n is calculated as the likelihood of choosing any lower item (e.g., *disagree*) as compared to the next higher item on the scale (e.g., somewhat *disagree)*. As the Likert scale of the GLRSS ranges from *strongly disagree* to *strongly agree,* there are 7 total categories, creating 6 category thresholds. Higher and positive threshold subscripts ($c_{4\text{-}6}$) indicate increased sensitivity in differentiating between more positive or higher values of the latent trait (θ), while lower or negative subscripts ($c_{1\text{-}3}$) reflect a decrease in sensitivity of the item to distinguish between categories.

There were several items that had very high sensitivity within the lower ranks of the category thresholds, indicating that the scale was exceptional at discovering when someone reported their relationship as unsatisfactory. Only one item was found useful at discovering relationships that are both dissatisfying and satisfying, Item 13, "*I often tell my partner that I love him/her*" (Table 1).

Chi-square diagnostics

The chi-square diagnostics are used in IRM to determine whether an item is functioning ("behaving") the way it is expected to. A non-significant chi-square indicates that there is no significant statistical difference between the expected and observed outcome of the item (scores) relative to the latent trait (θ). A significant chi-square means that there is a difference, and that the item is considered to be "misbehaving," i.e., not measuring the trait that it is supposed to. In the case of the Relationship Satisfaction subscale, only two items had significant chi-square results; Item 5, "*My partner and I have an active social life*," and Item 15, "*I always seem to be in the mood for sex when my partner is.*" This suggests these items have a poorer fit with the Relationship Satisfaction concept (θ).

Inspecting graphical output

Item information curves show the individual amount of variance and sensitivity (information gathered/shown) across all items on the scale. Figure 1

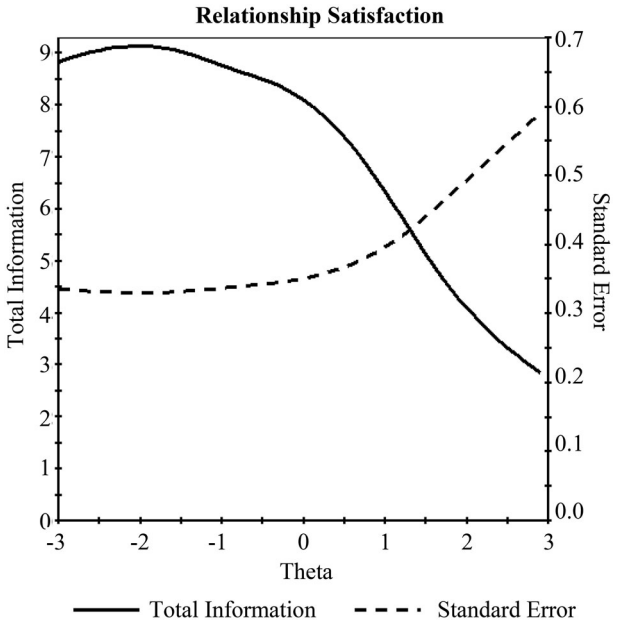

Figure 1. Total information curve–GLRSS relationship satisfaction.

displays the total information curve (solid line) for the Relationship Satisfaction subscale. The total information curve summarizes the ability of the subscale to provide information about the trait (θ), Relationship Satisfaction. Evaluation of the total information curve includes examining the standard error curve (broken line). Where the two curves meet is the value of θ where the level of error exceeds the amount of reliable information gathered. For the Relationship Satisfaction subscale, a theta level of approximately 1.33 is where the level of error begins to outweigh the amount of reliable information being gathered. This indicates that the Relationship Satisfaction subscale is better at distinguishing individuals who are in dissatisfying relationships (lower or negative values of θ).

Social support

Examining slope discriminants

Item-level parameter estimates for the Social Support subscale are presented in Table 2. When examining the slope discriminants for the Social Support subscale (A), only one item was found to be below acceptable levels. Item 23, "*I have told my co-workers about my sexual orientation/attraction,*" has a slope discriminant of .69, which falls short of acceptable. All of the other items are acceptable (above .75). This indicates that 7 of 8 items are capable of distinguishing the trait being measured (Social Support), while that one item has more difficulty.

Table 2. Item level parameter estimates and diagnostic statistics–GLRSS support.

Item		A	c_1	c_2	c_3	c_4	c_5	c_6	x^2	Probability
17	My family accepts my relationship with my partner	2.76	5.49	4.03	3.16	2.60	1.60	−.49	86.75	.0207
18	My partner's family accepts our relationship	1.04	2.91	2.50	1.81	1.20	.45	−.69	115.70	.0084*
19	My family would support our decision to adopt or have children	2.82	4.96	3.81	3.11	1.43	.85	−1.10	87.38	.0780
20	My partner's family would support our decision to adopt or have children	1.18	2.95	2.27	1.81	.67	.09	−1.21	141.46	.0001*
21	I feel as though my relationship is generally accepted by my friends	1.24	5.67	4.93	4.50	3.74	2.88	.74	45.56	.1861
22	I have a strong support system that accepts me as I am	1.28	6.36	5.62	4.66	3.47	2.26	.36	55.55	.1578
23	I have told my coworkers about my sexual orientation/attraction	.69	4.06	3.53	2.85	1.87	1.28	.01	82.23	.0855
24	Most of my family members know about my sexual orientation/attraction	.94	4.65	3.19	2.88	2.55	1.96	.47	64.42	.1154

*x^2 significant at $p < .01$.

Comparing category thresholds

The category thresholds for the Social Support subscale are distributed more evenly across items than those found in the Relationship Satisfaction subscale. As shown in Table 2, almost all items have positive indicators between levels, providing adequate sensitivity and ability to distinguish between trait levels. The first four items have a negative category threshold when distinguishing between the top-most scale indicators (*agree* vs. *strongly agree*); meaning that their ability to assess individuals with positive support systems is more difficult.

Chi-square diagnostics

There were two items with significant chi-square loadings. Chi-squares for Items 18 ("*My partner's family accepts our relationship*") and 20 ("*My partner's family would support our decision to adopt or have children*") were significant at the $p < .01$ level, indicating that these items were not "behaving" or measuring the trait in the way that was expected. In contrast, the chi-square tests for the parallel items (Items 17 and 19) about respondent's own family are not significant.

Inspecting graphical output

The information curves and standard error curves cross twice on the total information curve for Social Support (Figure 2). The crossing to the right in Figure 2 is at a theta trait of approximately 1.1. Again, the Social

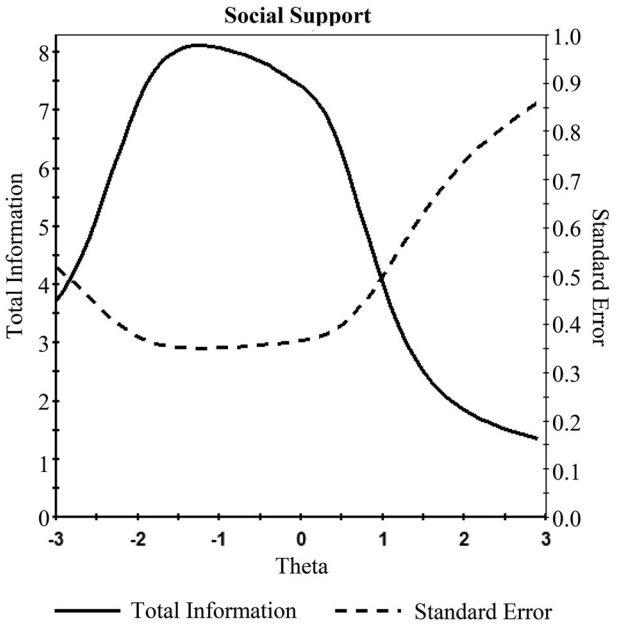

Figure 2. Total information curve–GLRSS social support.

Support subscale is helpful in providing information about the lack of social support, but not strong at uncovering high levels of support. The crossing to the left is very low on the graph, approximately a theta trait of −2.9, so low on the trait continuum that is not useful information.

Discussion

Relationship satisfaction

All items in the Relationship Satisfaction subscale had acceptable slope discriminants, indicating that they measured the underlying trait (θ). Two items were identified as "misbehaving" in the chi-square tests, i.e. items that do not measure and provide information in the way expected. Both items were modified from the Relationship Assessment Measure for Same Sex Couples, another scale of relationship satisfaction (Burgoyne, 2001). They were kept throughout the item reduction process in the CCT analyses and were retained because they fit the factor structure identified in those earlier analyses. However, they did carry information about levels of theta (Relationship Satisfaction) in the IRM, so their inclusion is warranted.

Examining all of the available IRM data, it is apparent that the Relationship Satisfaction subscale is better at distinguishing unsatisfying than satisfying relationships. As an individual reports greater satisfaction, the measure becomes less sensitive in distinguishing the amount of the trait (Relationship Satisfaction) that an individual is reporting. Several items are better at identifying unsatisfying relationships than others, for example, Item 1 "*My mate has the qualities I want in a partner*." This item has a high discriminant slope, appropriate levels of distinction between categories (information curve) and is "behaving" (chi-square test). Overall, the IRM analysis suggests that the Relationship Satisfaction subscale of the GLRSS is acceptable, in that it is able to distinguish among levels of the trait being measured.

Social support

In the Social Support subscale, one item failed to achieve an acceptable slope discriminant, suggesting that it was not measuring the latent trait of Social Support. Item 23, "*I have told my coworkers about my sexual orientation/attraction,*" was added based on a literature review suggesting that being "out" in the workplace led to higher levels of personal satisfaction and increased productivity. As the item response approach was not used as a data reduction technique in the previous study, the item stayed in the final version of the GLRSS. However, it is an item that requires further analysis. All other items in the scale had satisfactory slope discriminants, indicating that they were in fact measuring the trait *Social Support*.

However, two items about the partner's family had significant chi-square values, meaning they were not "behaving" the way they should. These items are theoretically, if not in a practical sense, related to social support in terms acceptance of the relationship and whether the partner's family would approve of the decision to have children. In contrast, the parallel questions about the respondent's family of origin support carry a great deal of information about how important the respondent's family's *lack* of support is to the respondent's report of low Social Support.

As was true of the Relationship Satisfaction subscale, the Social Support subscale is better at distinguishing among lower levels of social support than among higher levels of support. Further, it is evident from the information curves that the Social Support subscale is better at identifying lower levels of support than is the Relationship Satisfaction subscale at reporting higher levels of satisfaction (Figure 2 vs. Figure 1). The Social Support subscale item information graphs are also more distinct in describing a person's level of support and the ability of the scale to provide information on social support. The IRM suggests that it is, in fact, measuring the latent trait of Social Support.

Clinical applications

The GLRSS has several possible clinical applications. As previously noted, gay and lesbian populations often deal with higher levels of stress because of stigma from the majority population and social culture, and a lack of social support (Belous & Wampler, 2016). In general, this specific population seeks outpatient mental health services more frequently than other populations. It is more appropriate and ethically responsible to use an instrument that specifically addresses the culture of the people who are receiving therapy. Clinicians can obtain more specific information about the issues that someone in a gay or lesbian relationship may face by using the GLRSS. Furthermore, clinicians will be able to identify the stressors within the relationship with greater accuracy. As social support is such a key aspect in the lives of the gay and lesbian population, it is important the GLRSS is able to separate and measure social support as a separate construct from relationship satisfaction. Clinicians will be able to use this information to identify certain aspects of the relationship that may be causing issues and, therefore, better address them.

Research applications

The outcome of this paper using item response modeling to test the GLRSS provides more and significant evidence of the quality of this instrument.

Use of the GLRSS in research provides in-depth accurate information regarding the relationship satisfaction and social support of individuals in gay and relationships. Furthermore, using IRM as a litmus test when developing new instruments allows researchers to develop and test the items to ensure their quality and reliability. Using IRT/IRM to test the reliability and validity of any scale is important in determining the quality of each item as opposed to the instrument as a whole. This allows researchers to look more closely at each specific question and make adjustments rather than using trial and error of making adjustments to an instrument to improve the quality of the instrument as a whole.

Future developments

This study utilized a unidimensional graded response model and so could only test each subscale independently. A multidimensional model could provide more insight into the interactional elements between the subscales and the overall scale. Additionally, this analysis was conducted with previously collected data used for the development of the scale. A more stringent examination of that data would include samples from other populations. As the original study was successful in the development of the scale utilizing CTT, it stands to reason that we would find favorable results with this item analysis methodology–though that was not guaranteed. Additionally, the sample population was well-adjusted, typically, and was not reporting particularly distressing couple relationships. Thus, the data may be skewed toward a positive perspective overall. While this was a diverse sample from inside the United States, that was 100% of the sample and it was largely non-Hispanic whites. Translating the instrument into other languages or focusing on other ethnicities within the US might alter the results. We would also caution the use of this scale with bisexually identified persons in relationships, as the data that we have used for analysis in the original development, and this subsequent IRT-IRM validation study, did not include a large enough sample to be influential or distinct in the results, and so was omitted from analysis. Due to this, we do not have a way in which to determine if this measure adequately gathers data related to relationships in which one or more persons is bisexual in their sexual attraction identification. Being bisexually identified in a same gender (or different gender) relationship has unique and often difficult challenges other than that for those who are monosexually attracted (e.g. Gay or Lesbian identified) (Feinstein & Dyar, 2017).

Future studies with the Gay and Lesbian Relationship Satisfaction Scale need to be conducted with varied samples, for example, those that are in therapy for couple relationships, where both or either relationship

satisfaction and social support would be expected to have greater variation. Looking at differences in each partner's responses would also be important. Additionally, it would be judicious to look at relationships where one or more persons identify as bisexual gender non-conforming, fluid, or trans* to determine the ability of the scale to be used in such samples.

Conclusion

In sum, this secondary analysis of the data for the Gay and Lesbian Relationship Satisfaction Scale showed further evidence of the construct and content validity of the scale; the two subscales and the overall scale itself was shown to provide high levels of information on the trait being measured (relationship satisfaction and social support), and was able to distinguish between same gender relationships in which the respondent is satisfied, or dissatisfied in the relationship. Items were shown to provide high levels of trait characteristics, indicating that the items were good at providing and distinguishing between satisfied and unsatisfied couples, however the scale overall was better at identifying unhappy couples. This scale is appropriate for use with same gender couples in clinical and research settings, but can use additional research to expand the applicability to a wider cultural population.

Notes on contributors

Christopher K. Belous, PhD, is Associate Professor in the MS in Marriage and Family Therapy Program, Department of Behavioral Sciences, and Director of the Couple and Family Therapy Center, Purdue University Northwest.

Richard S. Wampler, PhD, Michigan State University (Retired), Dept. of Human Development and Family Studies, Couple and Family Therapy Doctoral Program. Texas Tech University (Emeritus), Dept. of Community, Family, and Addiction Sciences, Doctoral and Master's Programs in Couple and Family Therapy.

Britney L. Ledford, MFT, is Director of Youth Outpatient Services at Peachford Hospital in Dunwoody, GA.

References

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Waveland Press.

Baker, F. B. (2001). The basics of item response theory. Second Edition. Baltimore, MD: ERIC Clearinghouse.

Belous, C. K. (2015). Couple therapy with lesbian partners using an affirmative-contextual approach. *The American Journal of Family Therapy*, *43*(3), 269–281.

Belous, C. K., & Wampler, R. S. (2016). Development of the gay and lesbian relationship satisfaction scale. *Journal of Marital and Family Therapy*, *42*(3), 451–465.

Belous, C. K., Wampler, R. S., & Warmels-Herring, T. (2015). Gay male adaptation in the coming out process. *Journal of Gay & Lesbian Mental Health*, *19*(1), 55–71.

Burgoyne, R. W. (2001). The relationship assessment measure for same-sex couples (RAM-SSC): A standardized instrument for evaluating gay couple functioning. *Journal of Sex & Marital Therapy*, *27*(3), 279–287.

Busby, D. M., Christensen, C., Crane, D. R., & Larson, J. H. (1995). A revision of the dyadic adjustment scale for use with distressed and nondistressed couples: Construct hierarchy and multidimensional scales. *Journal of Marital and Family Therapy*, *21*(3), 289–308.

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). IRTPRO for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Carpenter, C., & Gates, G. J. (2008). Gay and lesbian partnership: Evidence from California. *Demography*, *45*, 573–590.

Cook, D. A., & Hatala, R. (2015). Got power? A systematic review of sample size adequacy in health professions education research. *Advances in Health Sciences Education*, *20*(1), 73–83.

Embretson, S. E., & Reise, S. P. (2000). Multivariate Applications Books Series. Item response theory for psychologists. Lawrence Erlbaum Associates Publishers.

Feinstein, B. A., & Dyar, C. (2017). Bisexuality, minority stress, and health. *Current Sexual Health Reports*, *9*(1), 42–49.

Gates, G. J. (2011). *How many people are lesbian, gay, bisexual and transgender?* The Williams Institute.

Hambleton, R. K., & Jones, R. W. (2005). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, *12*(3), 38–47.

Frost, D. M., & Meyer, I. H. (2009). Internalized homophobia and relationship quality among lesbians, gay men, and bisexuals. *Journal of Counseling Psychology*, *56*(1), 97–109.

Kuyper, L., & Fokkema, T. (2011). Minority stress and mental health among Dutch LGBs: Examination of differences between sex and sexual orientation. *Journal of Counseling Psychology*, *58*(2), 222–233.

Malyon, A. K. (1982). Psychotherapeutic implications of internalized homophobia in gay men. *Journal of Homosexuality*, *7*(2-3), 59–69.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, *129*(5), 674–697.

Nguyen, T. H., Han, H., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *The Patient–Patient-Centered Outcomes Research*, *7*(1), 23–35.

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded scores*. (Psychometric Monograph No. 17). Psychometric Society.

Samejima, F. (2016). Graded response models. In W. J. Van Der Linden (Ed.), *Handbook of item response theory* (pp. 95–108). Taylor & Francis.

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family*, *38*(1), 15–28.

Waring, E. M. (1984). The measurement of marital intimacy. *Journal of Marital and Family Therapy*, *10*(2), 185–192.

Waring, E., & Reddon, J. (1983). The measurement of intimacy in marriage: The waring intimacy questionnaire. *Journal of Clinical Psychology*, *39*(1), 53–57.

Yang, F. M., & Kao, S. T. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171–177.