DOI: https://doi.org/10.24297/jssr.v18i.9196

Exploring Reliability and Validity for the Professional Quality of Life Scale

Caitlyn C. McClure[1], Jennifer June Anderson[2]

[1]Northern Illinois Recovery Center

[2]University of Wisconsin-Whitewater

cmcclure@northernillinoisrecovery.com

Abstract

The Professional Quality of Life (ProQOL) instrument has been in use in some variation for decades with the most recent iteration becoming available in 2009. With twelve years of changes in the workforce plus applications of the ProQOL in novel professional groups, it is critical to determine if the tool is still reliable and valid. Data obtained from specific occupational groups, specifically college resident assistants (RAs) and substance abuse professionals (SAPs), will be utilized to explore if the measure holds up to past research claiming high reliability and validity. It is possible that the RAs (n=358) and the SAPs (n=40) will be outliers from these claims due to their unique positionality in their roles or due to the amount of time that has passed since the last update of the ProQOL.

Keywords*:* professional quality of life, instrument, reliability, validity

Introduction

Measuring the attitudes, beliefs, and wellness of a variety of professionals is a frequent practice nowadays (Groebe et al., 2018; Kodaka et al., 2011). As the "soft sciences" sought to demonstrate rigor and credibility among the scientific communities, they looked to quantify information beyond mere observation or description. The challenge with incorporating numerical measurement into the study of human behavior and experience is that such rating scales can be one-dimensional and miss the essential context required to gather the reason behind the numerical data collected. Because of the fluid and changing nature of humanity within the social context, it became critical for tools and instruments to prove that they could accomplish what was intended. Validity and reliability become issues of critical importance.

Validity occurs when an instrument accurately measures what it intends to, and reliability is demonstrated when the instrument consistently measures what it is supposed to. The difficulty with tools measuring beliefs, attitudes, feelings, or experiences of people is that these things change by their very nature. Further, beyond being variable day to day for a single individual person, sometimes there are larger cultural consideration, societal shifts, or trained patterns within populations that make an instrument that *was* valid and reliable into an obsolete tool. For these reasons, it is critical to continually test even highly utilized and praised instruments over time to confirm they still meet criteria to be statistically useful for their purposes.

The Professional Quality of Life (ProQOL) scale, the subject of this research study, is proof of the importance of continued adaptation to maintain relevance. Prior to the ProQOL as it is known today was the Compassion Fatigue Test developed by Charles Figley in the late 1980s (Stamm, 2002). Shortly after, he began collaborating with Stamm and they further explored the underlying constructs and created the Compassion Satisfaction and Fatigue Test in 1993 (Stamm, 2010). While their partnership was effective, they amicably agreed to shift ownership of the measure to Stamm in the late 1990s at which time it came under the title it holds today.

The Professional Quality of Life scale has been in use since 1995 and is now in its fifth official version (The ProQOL measure in English and non-English translations, 2019). It came on the heels of three different works published in the same year that focused on the negative effects on caregivers who provided care to traumatized individuals (Stamm, 2010). This was proof that a wave of interest in widening the traditionally narrow scope of trauma had begun. As the years have passed, interest has dramatically increased in this area of study with thousands of cases being contributed to the ProQOL data bank to open the door to awareness and change (Center for Victims of Torture, 2019).

The 30-question scale breaks into three subscales: compassion satisfaction, secondary traumatic stress, and burnout. There is no overall combined score as this takes away from the intention of the tool to assess and

understand the nuances affecting providers. The goal of continued development and variations is to better capture the distinct elements of each of the three subscales at present by building better reliability and validity.

To date, the ProQOL has been used extensively across many diverse types of service providers including nurses, child protection workers, substance abuse counselors, veterinary services, cancer patients, spoken-language interpreters, and government staff among others (Geoffrion, 2019; Hunsaker et al., 2015; Perkins & Sprang, 2013; Scotney et al., 2019; Khera et al., 2018; Mehus & Becher, 2016; Dang et al., 2015). Additionally, it has had several notable adaptations including the ProQOL-HCV (Armstrong et al., 2016), ProQOL-HIV (Duracinsky et al., 2012), and those made to meet specific cultural and linguistic needs while retaining the intention of the tool (Kim, & Choi, 2019; Köverová, 2018; Misouridou et al., 2020; Samson et al., 2016). The ProQOL has been translated into 26 different languages and is publicly available to further expand its use and the validity research (The ProQOL measure in English and non-English translations, 2019).

The ProQOL scale boasts a history of high reliability and validity scores that make it a tool frequently used to assess a variety of professionals. Reliability as defined by Rubin and Babbie (2017) is "a matter of whether a particular technique, applied repeatedly to the same object, would yield the same result each time" (p. 199). In this regard, the tool is recognized as holding up this expectation, though more recent nuances have arisen about specific populations benefiting from slight modifications to increase reliability (Keesler, 2020; Hemsworth et al., 2018). Similarly, validity is conventionally defined in research as "the extent to which an empirical measure reflects the *real meaning* of the concept under consideration" (Rubin & Babbie, 2017, p. 202). While the ProQOL tool has an extensive history of also meeting this criteria, more recent research for specific populations such as nurses suggests some revisions to certain questions under the subscales to improve validity (Hagan, 2019).

Additional research suggests that the current subscales could be individually improved especially with relation to the BO and CS elements, though the STS subscale is shown to already provide a valid scaling of unique individual differences beyond the general factor (Geoffrion et al., 2019). While recent research brings to light areas for improvement of the scale, it is because the current version of the ProQOL was released in 2009 and is due for some updates to maintain relevance to the world of professionals today. These updates encouraged by other authors do not discount the entirety of the ProQOL tool, but rather suggest adaptations to a changing workforce.

Methods

Through a strategic means comparison of each question of the ProQOL from hundreds of surveys across disciplines, the goal is to identify questions of concern that may disrupt the value of the subscales or the ProQOL tool at large. The quantitative data collected in the form of ProQOL responses from the previously completed research studies will allow for a thorough and robust data pool to be analyzed. Suggestions will be made about adaptations to the instrument that could improve its validity and reliability in more applications.

Participants

Due to the nature of this research and its focus on understanding the ProQOL as an instrument, no participants were directly elicited for this study. Rather, anonymized ProQOL scores from previous IRB-approved research efforts were aggregated. The Resident Assistant (RA) data comes from a small, rural, non-profit co-educational academic institution in the Midwest over a period of time from 2019-2020 (McClure et al., in press). The Substance Abuse Professional (SAP) data comes from participants recruited via addictions/social work professional group membership listservs and diverse interprofessional practice communities where substance abuse treatment and intervention were provided (Anderson et al., 2021).

Ethical Considerations

Due to the nature of the analysis being conducted on data collected under other IRB-approved studies, there is no identified risk to respondents. The datasets from the earlier studies have no identifiers and the anonymized information does not jeopardize the wellbeing or livelihood of respondents as a result. As a measure of certainty, an IRB (Institutional Review Board) packet was sent for an exemption review and was granted exemption.

Procedures

Because the immediate question of this study is about the ProQOL as a tool and is not focused specifically on how professionals and pseudo professionals are impacted differently across time, data from pre- and post-surveys in the RA study has been aggregated and analyzed together to build a more powerful dataset. To determine if there is any relevance in response patterns based on the type of provider responding to the survey, the survey data collected from the SAPs has been analyzed separately initially. These two separate analyses for RAs and SAPs allow for a clearer understanding of emergent patterns in less relevant questions of the ProQOL based on type of occupation. For further clarification and exploration, the scores of both RAs and SAPs have also been aggregated and then analyzed as a third dataset.

Measures

The ProQOL instrument consists of 30 questions: each with a 5-point Likert scale rating with 1 being never and 5 being very often. Participants are asked to view the questions in relation to their current work situation and select the number that reflects how frequently they have experienced the statement in the *last 30 days*. Higher average score indicates greater compassion satisfaction, burnout and secondary traumatic stress level. Within the ProQOL Secondary Trauma Stress Sub-Scale, several questions are inversely scored. Each sub-scale contains 10 specific questions and the same Likert scale rating. The ProQOL has three inherent sub-scales—one each for compassion satisfaction, burnout, and secondary trauma. The scale has a demonstrated history of high reliability scores and high construct validity (Stamm, 2010). The current version, the ProQOL 5, was used.

Data Analysis

Ensuring the reliability and validity of instrumentation is an essential element of determining the significance and real-world impact of results. Reliability is defined as having a tool that yields consistent results and this is further explored in ways such as test-retest reliability, inter-rater reliability, and internal consistency (Drost, 2011). Due to the nature of the ProQOL tool as a measurement of human attitudes and perceptions, test-retest reliability is not a helpful construct and inter-rater reliability does not apply in that the ProQOL is a self-report tool. For the purposes of this research, internal consistency has been utilized as the critical measure of reliability and will be determined by employing Cronbach's Alpha.

Cronbach's Alpha. Cronbach's Alpha is an accepted tool for determining internal consistency of measures as it assesses the degree of covariance of items in a scale relative to their sum score (Boateng et al., 2018). Cronbach's Alpha also offers a feature to demonstrate a corrected score output if specific items were removed entirely which allows for consideration in the present study of certain questions in subscales that jeopardize the internal consistency of each subscale. Mathematical research on this tool suggests the minimum sample size for it to be effective is 30 which allows for the analysis to be conducted by occupation as both the RA and SAP sample surpass this threshold (Bujang et al., 2018). There is also research cautioning against using Cronbach's Alpha independently to determine internal consistency and encouraging researchers to calculate it subscale-to-subscale rather than for the measure as a whole (Taber, 2018).

While reliability is one part of the determination of how the ProQOL functions, validity must also be assessed. Validity as a concept refers to how well an instrument measures what it is supposed to measure. This means that while a tool may be reliable, it may still be an overall faulty tool because it merely functions consistently but inaccurately. Focusing on certain areas of validity can prevent this situation.

Face validity merely asks that the content of the test appears on the surface to match its purposes. The ProQOL is a tool that has established face validity now and significant efforts in this writing will not be expended on confirming this type of validity. Content validity requires proof that a tool exhaustively represents all aspects of a construct which is not a claim of the ProQOL or a realistic expectation for an instrument measuring human experience and perception. As a function of the scope of this project, criterion validity will not be broached as it would be a more appropriate undertaking for a systematic review of various tools claiming to measure similar constructs. Thus, the most relevant type of validity research that has been undertaken in this project is construct validity and factor analysis will be the basis for grounding the supposition that the ProQOL exhibits construct validity.

Confirmatory Factor Analysis (CFA). Confirmatory factor analysis (CFA) is a multivariate procedure used to test how well the measured variables represent the number of constructs. In past research on this type of validity for the ProQOL tool, it has been suggested that certain questions better measure other constructs

than the ones they are assigned to at present (Sutjiono et al., 2019). CFA has the distinct benefit of determining specific items that contribute in a significant way to strong internal consistency and reliability which allows for improved understanding of potential shortcomings of strengths of the instrument (King et al., 2019). Ideally, this statistical analysis will be conducted on both RA and SAP scores separately and then aggregated to confirm that any patterns of reliability are translatable across professions.

Results

The data utilized in this analysis comes from two independent studies about RAs and SAPs. Each study utilized different expectations about gathering additional data about participants beyond the ProQOL scores therefore the descriptive statistical categories are not equivalent in both. In both studies, the participants are anonymous due to the intentional lack of descriptive data (RAs) or the nature of how the data was collected (SAPs).

Descriptive Statistics

RAs submitted 358 ProQOL surveys that could be utilized for instrument analysis. No further descriptive information is definitively known about this cohort of student employees. Certain assumptions can be made based on the nature of the role such as that RAs are generally undergraduate students with minimal direct training or education in the helping profession. Generally, RAs receive orientation and periodic training that may give specific strategies and feedback on managing individual and interpersonal difficulties.

SAPs submitted 40 surveys for further analysis and more descriptive information is available for this sample. Of the SAPS surveyed, 28 out of 37 (76%) have a master's degree or higher and 10 of 37 (27%) have earned some type of doctoral degree. 17 of 40 (43%) have more than 10 years of experience in the field and the majority report still providing direct support to consumers at 52.2%. All participants reported experiencing at least one unanticipated consumer death in the six months preceding the survey with 15% experiencing at least four consumer deaths in that same timeframe.

Some conclusions about these two groups can be made regarding descriptive statistics. It is reasonable to believe that the SAPs surveyed are older, more educated, and more experienced in their field and occupation while RAs tend to be younger, have less than an undergraduate degree, and may have little to no training in individual and interpersonal difficulties specifically. These stark differences are worth considering as the Cronbach's alpha scores are explored based on the current field of work.

Reliability

Due to the nature of the ProQOL as an instrument comprised of three, 10-question subscales this analysis has focused on the Cronbach's Alpha by subscale by occupational field. Cronbach's Alpha is a test used to determine the reliability of standardized items with any score above .7 being acceptable reliability, above .8 as ideal, and as close to 1.0 as possible (Salkind & Frey, 2020). For further review the Inter-Item Correlation Matrix was used to gain perspective about if items meant to measure the same subscale item were in fact doing so and demonstrating reliability. In these cases, a score of positive .3 or .4 is considered acceptable but closer to 1.0 is always better.

As seen in Table 1, the Cronbach's Alpha for subscales was .773 for burnout, .885 for compassion satisfaction, and .849 for secondary traumatic stress in the case of RAs. This indicates that acceptable reliability was achieved by the BO subscale while CS and STS reached a standard of high reliability. Meanwhile, with the SAPs, the Cronbach's Alpha for subscales was .716 for BO, .902 for CS, and .887 for STS. Notably, CS and STS are both above the ideal threshold and allow for the subscales to be considered reliable and in both samples the BO subscale is still in the acceptable range, but lower than its counterparts.

Table 1. Initial Cronbach's Alpha Scores

Reliability Statistics

Subscale by Group	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
RA BO	.773	.776	10
RA CS	.885	.886	10
RA STS	.849	.860	10
SAP BO	.716	.725	10
SAP CS	.902	.905	10
SAP STS	.887	.889	10

The Inter-Item Correlation Matrix and calculated scale variance if items were removed from the subscale indicated specific items of concern with regards to the subscales' reliability. As seen in Table 2, both RAs and SAPs were significantly impacted by the removal of question 15 (Cronbach's Alpha jumps from .773 to .792 and from .716 to .847 respectively) and question 29 (jumps from .773 to .779 and from .716 to .764 respectively) from the BO subscale. No concerns arose with the CS subscale.

Table 2. Burnout Cronbach's Alpha Impacts

Item-Total Statistics

Question Number	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
SAP Q15	22.53	25.025	-.645	.	.847
SAP Q29	22.35	20.644	-.205	.	.764
RA Q15	20.71	30.228	.187	.102	.792
RA Q 29	21.37	31.778	.193	.177	.779

Distinguished differences arose between the groups when considering these tests on the STS subscale as seen in Table 3. For RAs, questions 2 and 5 raised concern as the removal of each of these questions resulted in increased reliability of the scale. Removal of question 2 and 5 resulted in minor Cronbach's Alpha increase from .849 to .852 and from .849 to .853 respectively. SAPs had an impact on Cronbach's Alpha scores from the removal of question 28 with an increase from .887 to .904.

Table 3. Secondary Traumatic Stress Cronbach's Alpha Impacts

Item-Total Statistics

Question Number	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
RA Q2	17.74	38.357	.356	.183	.852
RA Q5	18.10	36.220	.404	.187	.853

SAP Q28	20.27	35.897	.187	.	.904

Validity

The 30 questions of the ProQOL were subjected to principal components analysis (PCA) to later explore the overall validity of the instrument. Prior to performing PCA, the suitability of data for the factor analysis was assessed based on criteria from Pallant (2020). Review of the correlation matrix showed numerous coefficients of .3 and above, making the next steps in analysis justifiable. Further, the Kaiser-Meyer-Olkin value was .903, significantly exceeding the recommended value of .6 (Kaiser, 1970, 1974). Bartlett's Test of Sphericity (1954) reached statistical significance in this sample supporting the factorability of the correlation matrix.

As reviewed in Table 4, PCA revealed the presence of six components with eigenvalues above one, explaining 25.5%, 16.6%, 6.0%, 4.6%, 4.0%, and 3.7% of the variance, respectively. A review of the scree plot revealed a major break after the second component and minor break after the third component, creating questions about the current subscale breakdown into three categories. However, through running a parallel analysis, three components emerged with eigenvalues exceeding the corresponding criterion values for a randomly generated data matrix of the same size (30 variables x 398 respondents). Due to the preexisting separation of the ProQOL questions into three subscales and the confirmation through parallel analysis that a three-component retention was appropriate, analysis proceeded with three components identified.

Table 4. PCA Eigenvalues and Variance

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings	
	Total	% of Variance	Cumulative %	Total	% of Variance
1	7.650	25.500	25.500	7.650	25.500
2	4.965	16.552	42.051	4.965	16.552
3	1.807	6.023	48.074	1.807	6.023
4	1.369	4.563	52.638		
5	1.197	3.989	56.627		
6	1.118	3.725	60.353		

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

As seen in Table 4, the three-component solution explained 48.1% of the variance with Component 1 contributing 25.5%, Component 2 contributing 16.6%, and Component 3 contributing 6.0%. To assist in the interpretation of these three components, oblimin rotation was performed revealing the presence of a simple structure (Thurstone, 1947) with components one and two having several strong loadings and the third component having noticeably fewer strong loadings. As seen in Table 5, there was no significant correlation between components one and two (r=.035) and components two and three (r=.191), however there was a low negative correlation between components one and three (r=-.412). The results of this analysis suggest some need for revision due to the third component, largely consisting of burnout subscale items (six out of seven items loading on component three are BO items), having overall less significant and weaker relationships, and indicating less validity.

Table 5. Component Correlation Matrix

Component	1	2	3	4	5	6
1	1.000	.035	-.412	.180	.075	-.232
2	.035	1.000	.191	-.398	-.156	-.093
3	-.412	.191	1.000	-.310	.016	.144
4	.180	-.398	-.310	1.000	.156	-.023
5	.075	-.156	.016	.156	1.000	.007
6	-.232	-.093	.144	-.023	.007	1.000

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization.

Verifying Factor Structure

Confirmatory Factor Analysis (CFA) was employed to verify the factor structure of a set of observed variables, in this case the items reliably measuring the subscale to which they belong. As described above, the data passed the initial tests to justify continuing with a CFA. The CFA was initially run with all 30 of the questions, then, in model two, with 26 of the 30 instrument items due to 4 items—two under CS and two under STS—being removed due to communalities scores lower than .5. In the third model, the problematic questions of 15 and 29 were removed rather than the others indicated followed by a fourth model removing both 15 and 29 as well as the previous four removed in model two. Model five came from a removal of all items with communalities scores below .6 followed by a final attempt to remove the same questions as five plus three additional items that had low loadings when 2 components were forced. Overall, as seen in Table 6, none of these models resulted in satisfactory results across multiple indices, a sign that the model itself may not have goodness-of-fit (Tabachnick & Fidell, 2019). Instead, as items were removed, while the Tucker-Lewis Index (TLI) and the Comparative Fit Index (CFI) improved but not into the acceptable range, the Root Mean Square Error of Approximation (RMSEA) continued to decline.

Table 6. CFA Results for 3 Component Models

Model	Chi-squared	Degrees of freedom	TLI	CFI	RMSEA
1	1285.959	392	.787	.821	.076
2	1038.245	286	.792	.831	.081
3	1100.050	344	.813	.842	.074

4	880.645	246	.818	.851	.081
5	498.578	100	.813	.863	.100
6	322.242	62	.833	.886	.103

Attempts to run the CFA as a two-component instrument rather than three were also unsuccessful. Through removal of questions that did not load significantly onto the two-component model, there were seven items left per component. On component one of this model, six of seven items were from the STS subscale. On the second component, all seven items were from the CS subscale. Running this two-component model that largely excludes BO as an alternative that is justifiable by earlier reliability and validity findings created no improvements in indices such as TLI, CFI, or RMSEA.

Conclusions

The goal of this statistical undertaking was to determine if the ProQOL was a valid and reliable instrument for measuring the burnout, secondary traumatic stress, and compassion satisfaction of individuals. More specifically, the question was if the ProQOL is truly as versatile as claimed in that it can be used on very different kinds of professionals and helpers with accurate results. Overall, questions were raised about the reliability of the burnout subscale on the instrument, a concern raised by other researchers while studying nurses (Hemsworth et al., 2018), and more specifically the reverse-coded items on the BO subscale, another concern raised in other literature (Sprang & Craig, 2015). With the two questions indicated as problematic by both RA and SAP data falling into the BO subscale, there is a possibility of universal concerns with items 15 and 29. The remaining items that were considered for removal to create slight improvements fell on the STS subscale and were different across occupations.

Apart from reliability concerns with the BO subscale, validity concerns also arose. There were significant issues with factor loadings on a three-component model in which six of seven items on that component were comprised of BO items and showed weaker and less significant relationships. Further, upon use of the CFA, no model versions constructed of the three-component or even two-component versions demonstrated true goodness-of-fit on indices including TLI, CFI, and RMSEA.

Limitations & Future Research

The statistical analyses outlined here have shed light on areas that were beginning to be brought out of the darkness. For instance, though the ProQOL is a popularly used research tool, there are some areas of concern related to certain subscales and potentially the structural underpinnings and assumptions. With literature pointing to other conceptualizations about the relationships between compassion satisfaction, burnout, and secondary traumatic stress in recent years, the next steps for research on the ProQOL instrument may involve some complex reconfiguring. It is possible that certain questions require linguistic changes to reflect more currently held understandings being explored, but also that certain subscales are inextricably linked and more work on parsimony may be necessary once informed by more recent theories about the subscale overlap areas. Further, limitations may need to be more clearly researched and outlined regarding the types of professionals that the ProQOL is most suited to.

Overall, the instrument has shown adaptability through the open format of using an anonymized data bank to explore these areas of strength and limitation. Seeing as the tool began with a singular focus on compassion fatigue and the researchers behind it have made changes as the understanding increases in this area of the field, it is likely that the ProQOL can be thoroughly assessed and shifted toward a more productive version of itself. Given the length of time since a new version has been released, some of the faulty elements are likely due to the natural changes in perspective and viewpoint since 2009. If the ProQOL can be reassessed with larger data sets from a variety of populations, the redevelopment and update of the tool can be informed by all the factors that could only briefly be reviewed in the scope of this work.

References

1. Anderson, J.J., Thomas, M.D., McClure, C., Winiecki, A., Lavin, M.M., Sorenson, B., & Gundacker, J. (2021). An exploratory study on the impact of the opioid epidemic on providers. *Traumatology.* Advance online publication. https://doi.org/10.1037/trm0000341

2. Armstrong, A.R., Herrmann, S.E., Chassany, O., Lalanne, C., Silva, M.H., Galano, E., Carrieri, P.M., Estellon, P., Sogni, P., & Duracinsky, M. (2016). The international development of PROQOL-HCV: An instrument to assess the health-related quality of life of patients treated for Hepatitis C virus. *BMC Infectious Diseases, 16*(1), 443. https://doi.org/10.1186/s12879-016-1771-0

3. Bartlett, M.S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16*(Series B), 296-298. https://doi.org/10.1111/j.2517-6161.1954.tb00174.x

4. Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H.R., & Young, S.L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*(N#), 149. https://doi.org/10.3389/fpubh.2018.00149

5. Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for Cronbach's Alpha Test: A simple guide for researchers. *The Malaysian Journal of Medical Sciences, 25*(6), 85–99. https://doi.org/10.21315/mjms2018.25.6.9

6. Center for Victims of Torture. (2019). *About us.* Professional Quality of Life Measure. https://www.proqol.org/About_us.html

7. Dang, W., Cheng, W., Ma, H., Lin. J., Wu, B., Ma, N., Wang, R., Xu, J., Zhou, T., & Yu, X. (2015). Reliability and validity of Professional Quality of Life Scale among government staff in earthquake-stricken areas in China. *Chinese Journal of Industrial Hygiene and Occupational Diseases, 33*(6), 440-443.

8. Drost, E.A. (2011). Validity and reliability in social science research. *Education Research and Perspectives, 38*(1), 105-124.

9. Duracinsky, M., Herrmann, S., Berzins, B., Armstrong, A.R., Kohli, R., Le Coeur, S., Diouf, A., Fournier, I., Schechter, M., & Chassany, O. (2012). The development of PROQOL-HIV: An international instrument to assess the health-related quality of life of persons living with HIV/AIDS. *Journal of Acquired Immune Deficiency Syndromes, 59*(5), 498-505. https://doi.org/10.1097/QAI.0b013e318245cafe

10. Geoffrion, S., Lamothe, J., Morizot, J., & Giguere, C-E. (2019). Construct validity of the Professional Quality of Life (ProQOL) scale in a sample of child protection workers. *Journal of Traumatic Stress, 32*(4), 566-576. https://doi.org/10.1002/jts.22410

11. Groebe, B., Strupp, J., Eisenmann, Y., Schmidt, H., Schlomann, A., Rietz, C., & Voltz, R. (2018). Measuring attitudes towards the dying process: A systematic review of tools. *Palliative Medicine*, *32*(4), 815–837. https://doi.org/10.1177/0269216317748889

12. Hagan, J. L. (2019). Psychometric evaluation of the ProQOL version 5 for assessing compassion satisfaction, burnout, and secondary traumatic stress in nurses. *International Journal of Studies in Nursing, 4*(3), 60-70. https://doi.org/10.20849/ijsn.v4i3.620

13. Hemsworth, D., Baregheh, A., Aoun, S., & Kazanjian, A. (2018). A critical enquiry into the psychometric properties of the professional quality of life scale (ProQol-5) instrument. *Applied Nursing Research, 39*, 81-88. https://doi.org/10.1016/j.apnr.2017.09.006

14. Hunsaker, S., Chen, H., Maughan, D., & Heaston, S. (2015). Factors that influence the development of compassion fatigue, burnout, and compassion satisfaction in emergency department nurses. *Journal of Nursing Scholarship, 47*(2), 186-194. https://doi.org/10.1111/jnu.12122

15. Kaiser, H. (1970). A second-generation Little Jiffy. *Psychometrika, 35*, 401-415. https://doi.org/10.1007/BF02291817

16. Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39*, 31-36. https://doi.org/10.1007/BF02291575

17. Keesler, J. M., & Fukui, S. (2020). Factor structure of the professional quality of life scale among direct support professionals: Factorial validity and scale reliability. *Journal of Intellectual Disability Research, 64*(9), 681-89. https://doi.org/10.1111/jir.12766

18. Khera, N., Ramos, G., Fruth, B., Arslan, W., Komenaka, I. K., Warsame, R. M., Northfelt, D. W., Griffin, J. M., & Sloan, J. A. (2018). Feasibility of a patient-reported outcomes quality-of-life instrument to improve care in underserved, ethnic minority patients with cancer. *Journal of Clinical Oncology, 36*(30), 65. https://doi.org/10.1200/JCO.2018.36.30_suppl.65

19. Kim, J. I., & Choi, H. S. (2019). Validation of Korean version of Professional Quality of Life (K-ProQOL) for nurses through confirmatory factor analysis. *Journal of the Korea Academia-Industrial Cooperation Society*, *20*(3), 243-251.

20. Kodaka, M., Poštuvan, V., Inagaki, M., & Yamada, M. (2011). A systematic review of scales that measure attitudes toward suicide. *International Journal of Social Psychiatry*, 57(4), 338-361. https://doi.org/10.1177/0020764009357399

21. Köverová, M. (2018). Konfirmačná faktorová analýza slovenskej verzie škály profesijnej kvality života (ProQOL) [A confirmatory factor analysis of the Slovak version of the Professional quality of life scale (ProQOL)]. *Ceskoslovenska Psychologie*, *62*(6), 565–578.

22. McClure, C., Anderson, J.J., Jang, E.Y., Birong, G., Brunn, P., & Krier-Jenkins, A. (in press). Understanding the impact of Resident Assistant experiences on their professional quality of life: The case for increased institutional supports. *The Journal of College and University Student Housing.*

23. Mehus, C. J., & Becher, E. H. (2016). Secondary traumatic stress, burnout, and compassion satisfaction in a sample of spoken-language interpreters. *Traumatology, 22*(4), 249-254. https://doi.org/10.1037/trm0000023

24. Misouridou, E., Pavlou, V., Kasidi, K., Apostolara, P., Parissopoulos, S., Mangoulia, P., & Fradelos, E. (2020). Translation and cultural adaptation of the Professional Quality of Life Scale (ProQOL V) for Greece. *Materia Socio-Medica, 32*(3), 187-190. https://doi.org/10.5455/msm.2020.32.187-190

25. Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Open University Press. https://doi.org/10.4324/9781003117445

26. Perkins, E. B., & Sprang, G. (2013). Results from the Pro-QOL-IV for substance abuse counselors working with offenders. *International Journal of Mental Health and Addiction*, *11*(2), 199-213. https://doi.org/10.1007/s11469-012-9412-3

27. The ProQOL measure in English and non-English translations. (2019). *Professional Quality of Life measure*. https://www.proqol.org/ProQol_Test.html

28. Rubin, A. & Babbie, E. R. (2017). *Research methods for social work* (9th ed.). Cengage Learning.

29. Salkind, N.J. & Frey, B.B. (2019). *Statistics for people who (think they) hate statistics* (7th ed.). SAGE Publications, Inc.

30. Samson, T., Lecovich, E., & Shvartzman, P. (2016). Psychometric characteristics of the Hebrew version of the Professional Quality-of-Life Scale. *Journal of Pain and Symptom Management, 52*(4), 575-581. https://doi.org/10.1016/j.jpainsymman.2016.03.019

31. Scotney, R. L, McLaughlin, D., & Keates, H. L. (2019). An investigation of the prevalence of compassion fatigue, compassion satisfaction and burnout in those working in animal-related occupations using the Professional Quality of Life (ProQoL) Scale. *The Veterinary Nurse, 10*(5), 276-284. https://doi.org/10.12968/vetn.2019.10.5.276

32. Sprang, G., & Craig, C. (2015). An inter-battery exploratory factor analysis of primary and secondary traumatic stress: determining a best practice approach to assessment. *Best Practices in Mental Health, 11*, 1–13.

33. Stamm, B. H. (2002). *Measuring compassion satisfaction as well as fatigue: Developmental history of the compassion satisfaction and fatigue test.* In C. R. Figley (Ed.), *Treating compassion fatigue* (pp. 107–119). Routledge.

34. Stamm, B. H. (2010). *The Concise ProQOL Manual* (2nd ed.). Pocatello, ID: ProQOL.org.

35. Taber, K.S. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education, 48,* 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

36. Tabachnick, B.G. & Fidell, L.S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

37. Thurstone, L.L. (1947). *Multiple factor analysis.* University of Chicago Press.

Conflict of Interest

There are no conflicts of interest to report.

Author Biography

Caitlyn McClure, DSW, LCSW currently serves as the Clinical Director at Northern Illinois Recovery Center focusing on meaningful program development and implementation for adults with substance use disorders. She has worked in the field of addictions and mental health in some capacity for over ten years with a focus particularly in trauma. She earned her Doctorate of Social Work in 2021 from Aurora University and took a deepening interest in employee burnout and secondary trauma through her research on the subject.