

Item-Level Analysis of the Revised Occupational Therapy Fieldwork Performance Evaluation Applied in Practice

Cynthia L. Sears¹, Brad E. Egan², Patricia F. Tomsic³, Craig A. Velozo⁴

¹Graduate College of Health Professions, Occupational Therapy Program, Hawai'i Pacific University, Honolulu, HI, USA

²College of Science and Health, Occupational Therapy Program, DePaul University, Chicago, IL, USA

³Department of Occupational Therapy, Wingate University, Wingate, NC, USA

⁴College of Health Professions, Department of Rehabilitation Sciences, Division of Occupational Therapy, Medical University of South Carolina, Charleston, SC, USA

Email: clsears@hpu.edu

How to cite this paper: Sears, C.L., Egan, B.E., Tomsic, P.F. and Velozo, C.A. (2024) Item-Level Analysis of the Revised Occupational Therapy Fieldwork Performance Evaluation Applied in Practice. *Open Journal of Therapy and Rehabilitation*, 12, 145-159. <https://doi.org/10.4236/ojtr.2024.122011>

Received: April 1, 2024

Accepted: May 21, 2024

Published: May 24, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Competency-based assessments for healthcare professionals are critical for safe and effective client outcomes. Rehabilitation clinical skill competency assessments must be validated and revised to produce safe and skilled practitioners. The revised American Occupational Therapy Association (AOTA) Fieldwork Performance Evaluation (FWPE) instrument measures occupational therapy student performance to determine readiness for practice. The assessment includes thirty-seven competencies that address both profession specific clinical skills and general professional behavior skills. The objective of this study was to use Rasch methods to explore the use of the revised FWPE in actual fieldwork practice and to determine the instrument's psychometric properties when separating the item components into two distinct subdomains: General Health Professions Competences and Occupational Therapy-Specific Competencies. Internal construct validity and test reliability were analyzed using data from 149 occupational therapy students after completing their initial Level II A fieldwork clinical internship. This study examined the item difficulty hierarchy, item fit, person-fit to model, person separation index, person separation reliability coefficient, strata, ceiling and floor effect, and unidimensionality of the FWPE instrument as a whole and as two separate domains. With the exception of not meeting the criteria for unidimensionality, the full FWPE instrument and the Occupational Therapy-Specific Competencies subdomain showed acceptable item-level psychometrics for reliability and precision. While the General Health Professions Competencies subdomain showed good item-level psychometrics, it was below the criterion for reliability and only separated the sample into

two strata. Results support the validity, reliability, and clinical use of the revised FWPE full instrument and the Occupational Therapy-Specific Competencies subdomain to measure entry-level clinical skill competencies in practice.

Keywords

Rehabilitation, Competency-Based Assessments, Occupational Therapy, Measurement, Rasch Analysis

1. Introduction

Assessing the clinical competency of pre-licensed healthcare providers is a major focus of academic preparation programs across many different professions. Academic programs are ultimately responsible for ensuring that graduates are ready and adequately prepared to enter practice [1] [2] [3]. Throughout didactic coursework, faculty may assess various components of clinical competency through direct observation, simulation, role play, and standardized patients. However, the evaluation of students' professional behaviors in real-world practice environments can be difficult in a classroom setting [4]. In many instances, it is not uncommon for professional academic accreditation standards to mandate the collection of students' professional and clinical competency outcomes during supervised practice-based learning [5]. Despite the clear necessity of practice-based learning opportunities, research across various health professions reveals numerous challenges and difficulties in assessing students' clinical competencies during rotations in authentic practice environments [1] [2] [6] [7] [8].

Several different factors have been identified as potentially challenging with respect to assessing students' professional skills and competencies. Dolan [1] suggested that there may always be a challenge to competency-based assessments because there continues to be no consensus on the definition of competency and judgments of competency are often subjective varying from practitioner to practitioner. To this end, Brown and colleagues [7] described the process of evaluating students' competency on fieldwork as abstract, noting that competencies may be understood differently in various clinical contexts and situations. Other studies also noted difficulties due to the increasing diversity of practice settings, differences in providing student supervision, changes in contemporary practice and service delivery models, and the potential lack of consistency among assessors [1] [6].

To address some of these concerns, one approach across rehabilitation professions has been to develop and validate standardized assessments for evaluating student competency [2] [3] [4] [5] [8]. Standardizing the assessment process can enhance objectivity. These assessments typically cover the full scope of the respective profession and include items that assess the knowledge, skills,

and behaviors expected of student clinicians prior to graduation. Moreover, they are often designed to assess students' competency over time and recognize that competency in certain skills may be more challenging to achieve than others [3] [4]. Best practices support regularly reviewing and updating competency-based tools to ensure that items remain relevant and effectively align with evolving professional practices and educational accreditation standards. In the United States, occupational therapy education programs have updated and modified standardized assessments to evaluate students' clinical competency on level II fieldwork since 1953 [9]. The most current tool, the revised American Occupational Therapy Association (AOTA) Fieldwork Performance Evaluation (FWPE), was revised and made available in 2020. Two versions of the revised FWPE exist, one for the occupational therapy assistant (OTA) student and another for the occupational therapy (OT) student. A task force appointed by AOTA informed the revisions to the original 2002 version of the tool. Preissner and colleagues [5] described the two-phase process of updating the assessment and the rationale for changing the rating scale and specific items to better reflect expanding fieldwork practice settings and current language in AOTA and accreditation documents. During the second phase, Kottorp and colleagues [10] conducted an item-level analysis of the revised FWPE using the Rasch measurement model by asking volunteer fieldwork educators to retrospectively complete a revised FWPE for a student they supervised and evaluated using the original FWPE four weeks prior. Although the psychometric findings suggested that the revised version of the FWPE demonstrated "good enough" unidimensionality given the variety of fieldwork experiences, the authors suggested that future investigations should continue to explore whether different subdomains within the scales might provide better unidimensional measures [10].

2. Purpose

This study aims to expand upon existing research in two distinct ways. It explores the revised FWPE's psychometric properties during actual supervised level II fieldwork experiences. Secondly, as suggested by Kottorp and colleagues, this study explored unidimensionality by dividing the tool into two discrete subdomains based on general professional behavior competencies and occupational therapy-specific competencies. These subdomains were based on previous studies that identified that many clinical competency student measures often include items relevant to general health professional behaviors, as well as, profession-specific skills [2] [3] [10].

3. Rasch Rating Scale Model

The FWPE data was analyzed using the Rasch Rating Scale model. The Rasch model places person, ability, and item difficulty measures on the same linear continuum. The model is based on the principle that a person with a higher

ability will have a greater probability of receiving higher ratings on a particular item than someone with a lower ability. Similarly, a person will be more likely to receive higher ratings on easier items than on more difficult items.

4. Methods

4.1. Instrument

The American Occupational Therapy Association's revised Fieldwork Performance Evaluation (FWPE) includes a rating and scoring guide, and the performance evaluation form includes six domains: Fundamentals of Practice, Basic Tenets of OT, Screening and Evaluation, Intervention, Management of OT Services, and Communication and Professional Behaviors. In total, 37 clinical competencies are assessed under the domains, and each is rated using a 4-point Likert scale to rate skill competencies ranging from 1-Unsatisfactory through 4-Exemplary Performance.

4.2. Study Design

This methodological study analyzed data collected from Level II A FWPE ($n = 149$) in 2021. The psychometric properties of the revised FWPE were evaluated using the Rasch model to assess item-level psychometrics, rating scale structure, dimensionality, model fit, person separation, and item-person difficulty/ability match.

4.3. Participants

This study utilized a convenience sample that consisted of 149 entry-level masters of OT students who participated in their first of two Level II fieldwork experiences during the spring 2021 and summer 2021 semesters. The students were located on two different campuses of the same institution. **Table 1** presents the demographics of the student participants. There were 75 student participants from campus A and 74 students from campus B. Fourteen participants were male, and 135 were female. There was limited diversity in the student population, which consisted of 88% White students and 12% Black, Asian, Hispanic, or Native American students. Student participants were placed in various clinical practice settings ranging from pediatric outpatient settings to inpatient acute care and skilled nursing settings. Adult physical rehabilitation settings accounted for 49% of the site settings, followed by 40% in pediatric settings and 11% of students in a school-based population setting.

4.4. Procedures

To investigate the item-level psychometrics of the revised FWPE, each item of the instrument was categorized into one of two conceptual subdomains, as seen in **Table 2**. After a thorough review and discussion of the items, three of the study's authors created a "General Health Professions Competencies" subdomain and an "Occupational Therapy-Specific Competencies" subdomain.

Table 1. Demographics of evaluated students.

Campus	A	75
	B	74
Gender	Male	14
	Female	135
Race	White	131
	Black	10
	Asian/Pacific Asia	3
	Hispanic	2
	Native American	3
Practice Settings	Early Intervention School	2
	Home Health	1
	Inpatient Acute Care	24
	Inpatient Rehabilitation	7
	Older Adult Community Living	1
	Outpatient Orthopedic	9
	Outpatient Rehabilitation	10
	Pediatric Community Based Practice	20
	Pediatric Outpatient in a Hospital	1
	Pediatric Outpatient Clinic	37
	School System	16
	Long Term Care	18
	Vocational Rehabilitation	2
	Other	1

Table 2. FWPE competency subdomains.

General Health Professions Competencies	Occupational Therapy-Specific Competencies
2 Adheres to safety	1 Adheres to the AOTA code of ethics
3 Prevent accidents	4 Articulates OT values
27 Demonstrate organ knowledge	5 Articulates value of occupation
28 Meets productivity standards	6 Articulates role OT
29 Communicates effectively	7 Rationale evaluation process
30 Produces documentation	8 Obtain evaluation information
31 Collaborates with educators	9 Selects screening/assessment
32 Seeks learning opportunities	10 Determines profile/performance
33 Responds to feedback	11 Analyzes client factors
34 Demonstrates proper workplace behavior	12 Administers assessments
35 Demonstrates time management	13 Modifies eval procedures
36 Manages relationships	14 Interprets eval results

Continued

37 Demonstrates respect for diversity	15 Documents eval results
	16 Articulates intervention
	17 Establishes client-centered
	18 Uses evidence to inform intervention
	19 Interventions that motivate and challenge
	20 Implement intervention plan
	21 Choses and modifies interventions
	22 Modifies for maximum potential
	23 Determines continuation of care
	24 Documents client's response
	25 Collaborates with an OT assistant/aide
	26 Understands OT costs and funding

General Health Professions Competencies were defined as common expectations for all health professions and do not require unique OT knowledge based on the item specification. Occupational Therapy-Specific Competencies were defined as the tenets of occupational therapy along with tasks associated with OT screening, evaluation, and intervention planning and processes, including any item where OT is explicitly named in the item specification. With these descriptions in mind, each of the three authors separately categorized the individual FWPE items into the created subdomains. Then, each author's list was compared and discussed as a group to reach a consensus in placing items under "General Health Professions Competencies" and "Occupational Therapy-Specific Competencies." Furthermore, a scatter plot of General Health Professions Competencies measures and Occupational Therapy-Specific Competencies measures was generated to determine the relationship between these measures, *i.e.*, (the percent of subdomain measures within measurement error and outside of the measurement error).

FWPE data was aggregated from the student participants during the 2021 calendar year. An administrative assistant created a Microsoft Excel spreadsheet per the parameters required for the Rasch study and deidentified participant information before sharing it with researchers. Next, the FWPE data was analyzed using Winsteps software (version 3.93.1; John Linacre/Winsteps.com, Beaverton, OR). This software provides instrument and item-level psychometrics and graphs to evaluate rating scale structure, dimensionality, model fit, person separation, and item-person difficulty/ability match. This research received IRB approval from a university that three of the authors were previously affiliated with.

4.5. Statistical Analysis

The rating scale structure was evaluated based on Linacre's three critical

rating-scale criteria: 1) each rating-scale category has a minimum of 10 observations, 2) average measures within each rating-scale category advance monotonically (*i.e.*, higher ratings will have higher average measures) and 3) outfit mean-squares for each rating-scale category are less than 2 [11].

The Rasch model assumes that the instrument is unidimensional, meaning the items of the instrument measure only one domain/construct. Principal components analysis (PCA) of residuals was used to evaluate the unidimensionality of the revised FWPE. PCA criteria for unidimensionality was based on Smith's [12] recommendations: 1) the Rasch dimension explains at least 40 percent of the variance, 2) the first contrast explains less than 4 percent of the variance, and 3) the eigenvalue of the first contrast is less than or equal to 2. If the revised FWPE did not meet the above three criteria, we investigated attenuated correlations (correlations corrected for measurement error) across three clusters of item loadings (high positive loadings, middle-level loadings, and low negative loadings). Attenuated correlations close to unity suggest that the clusters are highly correlated, and the instrument is essentially unidimensional Joreskog [13]. Linacre [14] suggests an acceptable criterion of attenuated correlations of 0.81 or greater.

Item and person fit represents how closely the item and person response patterns match those predicted by the Rasch model. Items and persons were classified as fitting the Rasch model if mean-square standardized residuals were less than 1.4 and standardized z-scores were less than 2 [15].

The reliability of the FWPE was evaluated using person separation, person separation reliability, and strata. Strata represent the number of "center points" in the distribution of person measures that are three standard errors apart, indicating the number of statistically distinct groups the instrument separates within the sample. The more strata, the more precise the instrument. The person separation index is expected to be 2 or greater, person separation reliability equal to or greater than 0.80, and strata 2 or greater [16] [17]. The following formula was used to calculate strata [15].

$$\text{Strata} = [4 * (\text{person separation index}) + 1] / 3$$

An item-difficulty hierarchy was used to evaluate the construct validity of the FWPE. Items with low calibrations (easy items) represent items expected to be easily accomplished by students, and items with high calibrations (hard items) represent items expected to be difficult for students. Since the Rasch model places items and persons on the same linear continuum, the relative difficulty of the FWPE for the sample can be determined. If the FWPE were easy for the sample, we expect the person mean to be higher than the item mean. Ceiling or floor effects were determined if 15% or more individuals received maximum or minimum measures, respectively [18].

5. Results

Table 3 presents the summary of the item-level psychometrics of the instrument.

Table 3. Psychometric analysis results.

Analyses	Full Instrument	General Health Professions Competencies	Occupational Therapy-Specific Competencies
	(37 items)	(13 items)	(24 items)
N	149	149	149
Principal Components Analysis	Met 1/3 criteria	Met 0/3 criteria	Met 0/3 criteria
Disattenuated Correlations	2/3 above 0.80	2/3 above 0.80	3/3 above 0.80
Rating Scale	Met 2/3 criteria	Met 2/3 criteria	Met 2/3 criteria
Item Fit to Model	95% (35/37)	100% (13/13)	96% (23/24)
Person Fit to Model	92% (137/149)	97% (144/149)	95% (142/149)
Person Separation Index	3.28	1.69	2.75
Person Separation Reliability	0.91	0.74	0.88
Strata	4.71	2.59	3.08
Ceiling/Floor	2% max (3/149)/0% min	15% max (23/149)/0% min	3% max (4/149)/0% min

In general, item-level psychometrics show mixed results across the full instrument. Regarding dimensionality, the instrument failed to demonstrate unidimensionality, meeting 1/3 PCA criteria. Since the instrument failed to meet PCA criteria, we investigated the disattenuated correlations across high-positive, middle, and high-negative loading clusters. The instrument had 1/3 cluster correlations failing to meet the criterion.

Regarding rating scales, the full instrument and subdomain measures met 2 of 3 criteria; the lowest rating scale was not used across all items. The potential influence of scoring instructions on this finding is presented in the discussion.

Items and person fit the Rasch measurement model and were good across the full instrument and subdomains, ranging between 92 - 100 percent, with the best item and person fit demonstrated by the General Health Professions Competencies subdomain. The full instrument showed the best overall measurement qualities, including person separation, person separation reliability, and strata, separating the sample into four statistically distinct strata. The Occupational Therapy-Specific Competencies subdomain showed the next best separation statistics, separating the sample into three statistically distinct strata. In contrast, the General Health Professions Competencies failed to meet the person separation criteria of 2 and the person separation reliability criteria of 0.80, only separating the sample into 2.59 statistically distinct strata. Only the General Health Professions Competencies measure showed a ceiling effect, with 15% of the sample receiving maximum measures.

Figure 1 is a scatterplot of Occupational Therapy-Specific Competencies against General Health Professions Competencies. Error bands are two standard errors above and below the mean. Symbols falling within the error bands indicate the students' occupational therapy-specific competencies and general

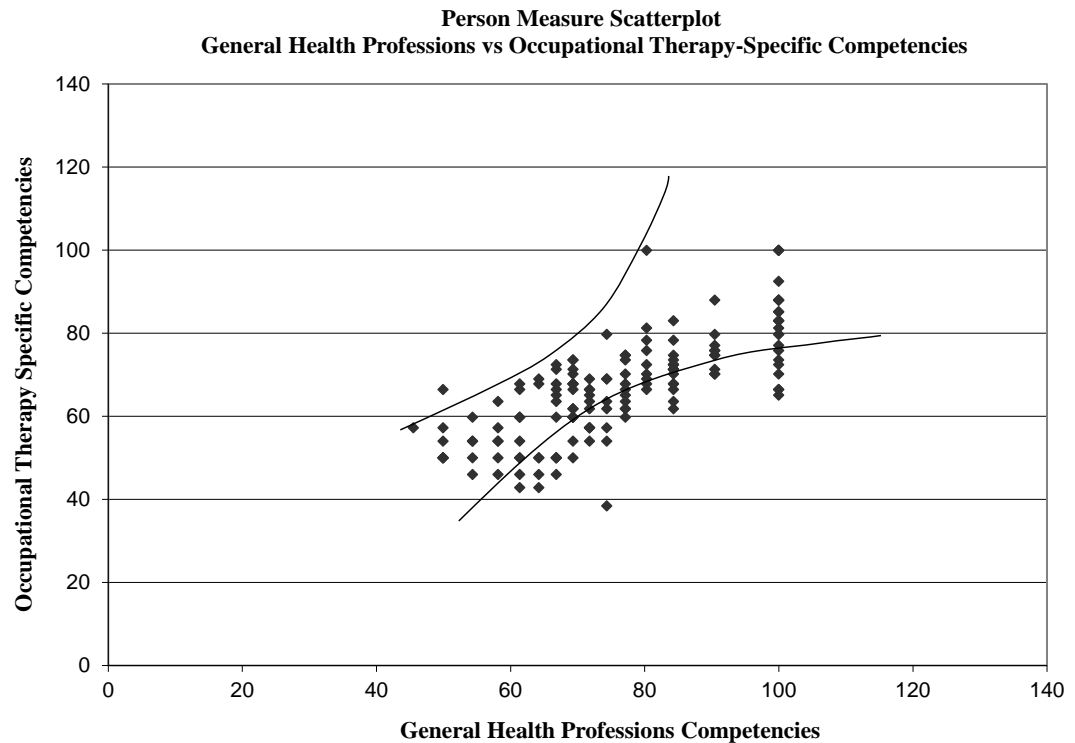


Figure 1. Person measure scatterplot.

health professions competencies were statistically equivalent. Symbols falling above the top error band indicate a student's measure was significantly higher for Occupational Therapy-Specific Competencies than for General Health Professions Competencies, and symbols falling below the bottom error band indicate a student's measure was significantly higher for General Health Professions Competencies than for Occupational Therapy-Specific Competencies.

Figure 2 presents a Wright map of the full instrument, placing person and item measures on the same scale. The original logit scale generated from Rasch analysis was converted to a 150-point scale to be similar to the raw-score scale of the FWPE. Student measures to the left of the vertical dashed line with the lowest ability students at 111 units and the highest ability students at 148 units. The student measures form a relatively normal distribution with the mean (M to the left of the vertical dashed line) at 127 units. Item difficulty measures are on the right of the dashed vertical line with the easiest items, "33 Responds constructively to feedback in a timely manner" and "34 Demonstrates consistent and acceptable work behaviors" at 102 units and the hardest item, "26 Demonstrates through practice or discussion an understanding of costs and funding systems related to occupational therapy services, such as federal, state, third party, and private payers" at 120 units with a mean (M to the right of the vertical dashed line) at 111 units. Of note is the number of items at the same measure, especially at 114 units and 117 units. In general, students are performing well on the revised FWPE (student measure mean of 127 units vs item difficulty measure mean of 111 units).



Figure 2. Wright map of full instrument.

Bolded/shaded items represent General Health Professions Competencies items, and non-bolded/non-shaded items represent Occupational Therapy-Specific Competencies items. The General Health Professions Competencies items are generally easier than the Occupational Therapy-Specific Competencies items.

Figure 2 presents a Wright map of student measures on the full revised FWPE (left of the dashed vertical line) and average item difficulties (right of the dashed vertical line). The full tool can be accessed and reviewed at <https://www.aota.org/-/media/corporate/files/educationcareers/fieldwork/fieldwork-performance-evaluation-occupational-therapy-student.pdf>. Shaded items represent the General Health Professions Competencies subdomain, and non-shaded items represent the Occupational Therapy-Specific Competencies. “.” represents the measure of 1 student, and “#” represents the measures of 2 students. “M” indicates mean, “S” indicates two standard errors, and T indicates three standard errors, with the letters to the left of the dashed vertical line referring to mean and standard errors of student measures and letters to the right of the dashed vertical line refer to mean and standard errors of item measures. “/” symbols indicate a break in the graph scale (i.e., between 0 and 100). Abbreviations: 1 AOTA ethics = 1 AOTA Code of Ethics, = 2 Adhere safety = 2 Adheres to safety regulations, 3 Prevent accident = 3 Preventing accidents, 4 Artic OT values = 4 Articulates OT values, 5 Artic value occupa = 5 Articulates value of occupation, 6 Artic role OT = 6 Articulates role of OT, 7 Rationale evaluation = 7

Rationale for evaluation process, 8 Obtain eval info = 8 Obtains information for evaluation, 9 Select screen/assess tools = 9 Selects screening/assessment tools, 10 Determ profile/perform = 10 Determines profile and performance, 11 Analyze client factors/context = 11 Analyzes client factors/context, 12 Admin assess = 12 Administers assessments, 13 Mod proced = 13 Modifies evaluation procedures, 14 Interp results = 14 Interprets evaluation results, 15 Doc results = 15 Documents results of evaluation, 16 Artic interven ration = 16 Articulates rationale for intervention, 17 Estab client-center plan = 17 Establishes client-centered plan, 18 Uses evid interven = 18 Uses evidence to inform intervention, 19 Interv motiv/challenge = 19 Intervention motivates/challenges client, 20 Implement intervent plan = 20 Implement intervention plan, 21 Choose/modif interven = 21 Chooses/modifies intervention, 22 Mod interv max perform = 22 Modifies to maximize performance, 23 Determ contin/discont = 23 Determines continuation/discontinuation, 24 Doc client response = 24 Documents client's response, 25 Collab assist/aide = 25 Collaborate with assistant/aide/etc., 26 Understand costs/fund = 26 Understands costs and funding, 27 Demon organization knowledge = 27 Demonstrates organizational knowledge, 28 Meets prod standards = 28 Meets productivity standards, 29 Communicate effective = 29 Communicates effectively, 30 Prod doc = 30 Produces documentation, 31 Collab FW educator = 31 Collaborates with FW educator, 32 Seek learn oport = 32 Seeks out learning opportunities, 33 Respond feedback = 33 Responds to feedback, 34 Demon workplace behavior = 34 Demonstrates acceptable workplace behavior, 35 Demon time manage = 35 Demonstrates time management, 36 Manage relation = 36 Manages relationships, 37 Demon respect diversity = 37 Demonstrates respect of diversity.

6. Discussion

Except for not meeting the criteria for unidimensionality, the full revised FWPE instrument and the Occupational Therapy-Specific Competencies subdomain show acceptable item-level psychometrics, reliability, and precision when used clinically. While the General Health Professions Competencies subdomain showed good item-level psychometrics, it was below the criterion for reliability and only separated the sample into two strata.

All instruments failed to meet all criteria for demonstrating unidimensionality. Kottorp and colleagues [10], found similar challenges in the full instrument meeting unidimensionality and suggested that the challenge in meeting unidimensionality may be a function of fieldwork performance addressing a wide variety of competencies. It should be noted that though the Occupational Therapy-Specific Competencies subdomain failed to meet any unidimensionality criteria, high attenuated correlations suggest that the items of this subdomain pragmatically work as a single instrument. Future studies should investigate further grouping items to determine if improved unidimensionality can be achieved.

The full instrument and subdomains met all criteria for the rating scale, except for no students receiving the lowest rating. Similarly, Kottorp and colleagues [10], found only 23 lowest ratings. This is likely a function of students being well-trained before participating in fieldwork and fieldwork educators avoiding rating students using the lowest rating on any item since this could result in the student failing the fieldwork experience.

Item fit to the Rasch model was good across the full instrument and subdomains, with only two items misfitting in the full instrument (#28 productivity and #35 time management); Kottorp and colleagues [10] found five items misfitting the Rasch model. This may have been a function of the responders in the Kottorp study basing their ratings on the memory of the student's performance rather than the more immediate ratings of using the revised FWPE during the student's actual placement. It should be noted that a single item, FWPE item #13, "Modifies evaluation procedures based on client factors and contexts," misfit for the Occupational Therapy-Specific Competencies subdomain. Similarly, 92% of persons fit the Rasch model for the present study and the Kottorp and colleague study [10].

The full instrument and Occupational Therapy-Specific Competencies subdomain's reliability and precision indicators were acceptable, with the full instrument showing the best precision. Our indicator of person separation of 3.28 was comparable to that of Kottorp and colleagues at 3.48 [10].

The item-difficulty hierarchy of the full instrument and subdomains was logical and has implications for the educational preparation of students for clinical practice. The findings show that the General Health Professions Competencies subdomain has a ceiling effect and students tend to get higher scores on the General Health Professions Competencies items than the Occupational Therapy-Specific Competencies items. The finding that General Health Professions Competencies items tend to be easier than the Occupational Therapy-Specific Competencies items suggest that there may be a natural progression in the development of clinical skills from general-to-specific competencies. The item-difficulty hierarchy demonstrated in this study replicates that found in Kottorp and colleagues [10]. In both studies, General Health Professions Competency items such as responding to feedback and demonstrating workplace behaviors are easier than Occupational Therapy-Specific knowledge and skills, such as understanding costs and funding and administering assessments.

The item-difficulty-hierarchical structure of the revised FWPE may also be the basis for personalized measurement, as Velozo [19] suggested in his 2020 Eleanor Clark Slagle lecture. The present findings form the basis for developing keyform ability maps whereby students can monitor their progress from general health professional skills to occupational therapy-specific skills. Sears and Egan [20] [21] have demonstrated the application of keyform ability maps as a signature instructional strategy within entry-level occupational therapy education. A General Health Professions Competencies item hierarchy key form may also be found to be useful across other rehabilitation professions and during

interprofessional experiences as a person-centered feedback tool or performance rubric.

7. Limitations

The current study had limitations that need to be considered in relation to the findings. The present study was limited to a convenience sample of students from two cohorts within one University's entry-level occupational therapy program. Moreover, the analysis only considered using final revised FWPE scores. Future studies may consider analyzing midterm scores along with final scores to more fully understand how the performance of certain competencies may change or remain stable over time. Lastly, this study references the revised FWPE for occupational therapy assistant students due to its shared items with the FWPE for occupational therapy students. However, only occupational therapy students were included in this study. Thus, there may be value in future studies that look exclusively at or include occupational therapy assistant students' fieldwork performance.

8. Conclusions

To ensure safe and effective client outcomes, occupational therapy students must be assessed for and prepared for readiness to enter clinical practice. The instrument's psychometric properties used to measure student skill competence are vital to ensure an ethical and equitable process for progressing students to independent practice.

In summary, the findings of this study support using the revised FWPE in clinical practice. While failing to meet unidimensionality criteria, the full instrument and Occupational Therapy-Specific Competencies subdomain showed good item-level psychometrics and precision. Furthermore, the item-difficulty structure of the revised FWPE may support personalized student self-monitoring of their progression in acquiring clinical skills.

The findings from this research contribute to the present literature supporting the assessment of the psychometric properties of high-stakes instruments used to measure competency in clinical skills for rehabilitation professionals entering into practice.

Acknowledgements

Caity Stratmeyer for administrative support, Occupational Therapy clinical educators' support of OT student success.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Dolan, G. (2003) Assessing Student Nurse Clinical Competency: Will We Ever Get

- It Right? *Journal of Clinical Nursing*, **12**, 132-141.
<https://doi.org/10.1046/j.1365-2702.2003.00665.x>
- [2] Martiane-Ramirez, N.L., Pineda-Galan, C., Rodriguez-Bailon, M. and Romero-Galisteo, R.P. (2022) Competence Assessment Rubric in the Physiotherapy Practicum. *PLOS ONE*, **17**, e0264120.
<https://doi.org/10.1371/journal.pone.0264120>
- [3] McAllister, S., Lincoln, M., Ferguson, A. and McAllister, L. (2011) A Systematic Program of Research Regarding the Assessment of Speech-Language Pathology Competencies. *International Journal of Speech-Language Pathology*, **13**, 469-479.
<https://doi.org/10.3109/17549507.2011.580782>
- [4] Judd, B.K., Scanlan, J.N., Alison, J.A., Waters, D. and Gordon, C.J. (2016) The Validity of a Professional Competence Tool for Physiotherapy Students in Simulation-Based Clinical Education: A Rasch Analysis. *BMC Medical Education*, **16**, Article No. 196. <https://doi.org/10.1186/s12909-016-0718-x>
- [5] Preissner, K., Duke, K.B., Killian, C., Ouyang, R.L., Jarek, E.D. and Kottorp, A. (2020) The Revised American Occupational Therapy Association Fieldwork Performance Evaluations: Evaluation of Content Validity-Part 1. *The American Journal of Occupational Therapy*, **74**, 7406205090p1-7406205090p13.
<https://doi.org/10.5014/ajot.2020.044180>
- [6] Allison, H. and Turpin, M.J. (2004) Development of the Student Placement Evaluation Form: A Tool for Assessing Student Fieldwork Performance. *Australian Occupational Therapy Journal*, **51**, 1440-1630.
<https://doi.org/10.1111/j.1440-1630.2004.00414.x>
- [7] Brown, T., Yu, M.L., Hewitt, A. and Etherington, J. (2020) Professionalism as a Predictor of Fieldwork Performance in Undergraduate Occupational Therapy Students: An Exploratory Study. *Occupational Therapy in Health Care*, **34**, 130-154.
<https://doi.org/10.1080/07380577.2020.1737896>
- [8] Rodger, S., Chien, C.W., Turpin, M., Copley, J., Coleman, A., Brown, T. and Caine, A.M. (2016) Establishing the Validity and Reliability of the Student Practice Evaluation Form-Revised (SPEF-R) in Occupational Therapy Practice Education: A Rasch Analysis. *Evaluation & the Health Professions*, **39**, 33-48.
<https://doi.org/10.1177/0163278713511456>
- [9] Crist, P.A. and Cooper, R.G. (1988) Evaluating Clinical Competence with the New Fieldwork Evaluation. *The American Journal of Occupational Therapy*, **42**, 771-773.
<https://doi.org/10.5014/ajot.42.12.771>
- [10] Kottorp, A., Killian, C., Duke, K., Leggett, C., Drasga, R. and Preissner, K. (2023) The Revised American Occupational Therapy Association Fieldwork Performance Evaluations: Evaluation of Internal Structure, Response Processes, and Precision-Part 2. *The American Journal of Occupational Therapy*, **77**, Article ID: 7705205060.
<https://doi.org/10.5014/ajot.2023.050207>
- [11] Linacre, J. (2002) What Do Infit and Outfit, Mean-Square and Standardized Mean? *Rasch Measurement Transactions*, **16**, 878.
- [12] Smith Jr., E.V. (2002) Detecting and Evaluating the Impact of Multidimensionality Using Item Fit Statistics and Principal Component Analysis of Residuals. *Journal of Applied Measurement*, **3**, 205-231.
http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=retrieve&db=pubmed&dopt=citation&list_uids=12011501
- [13] Joreskog, K.G. (1971) Statistical Analysis of Sets of Congeneric Tests. *Psychometrika*, **36**, 109-133. <https://doi.org/10.1007/BF02291393>

-
- [14] Linacre, J. (2007) Standard Errors and Reliabilities: Rasch and Raw Score. *Rasch Measurement Transactions*, **20**, 1.
- [15] Wright, B.D. and Masters, G.N. (1982) Rating Scale Analysis. Mesa Press, San Diego.
- [16] Linacre, J.W.C. (2012) WINSTEPS Rasch-Model Computer Programs in Version 5.6.0.0. Winsteps.Com, Beaverton.
- [17] Fisher Jr., W.P. (1992) Reliability, Separation, Strata Statistics. *Measurement Transactions*, **6**, 238. <https://www.rasch.org/rmt/rmt63i.htm>
- [18] Lim, C.R., Harris, K., Dawson, J., Beard, D.J., Fitzpatrick, R. and Price, A.J. (2015) Floor and Ceiling Effects in the OHS: An Analysis of the NHS PROMs Data Set. *BMJ Open*, **5**, e007765. <https://doi.org/10.1136/bmjopen-2015-007765>
- [19] Velozo, C.A. (2021) Using Measurement to Highlight Occupational Therapy's Distinct Value. *The American Journal of Occupational Therapy*, **75**, Article ID: 7506150005. <https://doi.org/10.5014/ajot.2021.746001>
- [20] Sears, C.L. and Egan, B.E. (2022) A Qualitative Study on Occupational Therapy Students' Perceptions of Using Keyform Maps. *Journal of Education and Training Studies*, **10**, 29-34. <https://doi.org/10.11114/jets.v10i2.5430>
- [21] Sears, C.L. and Egan, B.E. (2023) Promoting Person-Centered Measurement Using Keyform Maps within an Occupational Therapy Curriculum. *Journal of Instructional Pedagogies*, **29**, 1-13. <https://doi.org/10.11114/jets.v10i2.5430>