Psychometric Properties and Informant Agreement of the WHODAS 2.0 in Youth with Mental Disorder and their Parents

by

Erica McDonald

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Public Health and Health Systems

Waterloo, Ontario, Canada, 2021

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis is the work of Erica McDonald with the collaboration of her supervisor, Dr. Mark Ferro. Appropriate actions were taken to ensure confidentiality of the participants in the *MY Study* dataset.

Abstract

Background: Mental disorders affect one in five youth and the impact of these disorders extend across the life course. Functional impairment moderates many of the negative effects of mental disorders; however, its assessment can be elusive in the context of child psychiatry. Patient-reported data are integral for clinicians to obtain health information and evaluate the health and functioning of children with mental disorder. While there is value in obtaining health assessments from multiple informants, youth self-report and parent proxy-reports often disagree. This disagreement can be complex to manage and can negatively impact care and health outcomes for youth.

Objectives: This study estimated convergent/divergent validity, internal consistency, parent-youth disagreement, and factors associated with disagreement on the 12-item and 36-item versions of the World Health Organization Disability Assessment Schedule (WHODAS) 2.0, an assessment of functional impairment.

Methods: Data came from a sample of 56 youth aged 14-17 years with a common mental disorder who received either inpatient or outpatient mental health services at tertiary pediatric care center in Ontario, Canada. Correlations between the WHODAS 2.0 and domain scores on the KIDSCREEN-27, a health-related quality of life assessment, were used to assess convergent validity. Correlations between the WHODAS 2.0 and demographic variables (youth age, youth sex, and household income) were used to assess divergent validity. Internal consistency was measured using ordinal α. The Bland-Altman method and intraclass correlation coefficients (ICC) were used to assess parent-youth disagreement. Finally, logistic regression models were created to explore factors associated with clinically meaningful disagreement between parents and youth on the 36-item and 12-item version of the WHODAS 2.0.

Results: Correlations between WHODAS 2.0 scores and KIDSCREEN domain scores were low to moderate for both parents (τ= -0.42 to -0.05) and youth (τ= -0.41 to 0.01). Correlations between WHODAS 2.0 scores and demographic variables were low to moderate for both parents (τ/Point Biserial = -0.12 to 0.29) and youth (τ/Point Biserial = -0.06 to 0.32). All ordinal α values were >0.7. Therefore, internal consistency of both versions of the WHODAS 2.0 was

sufficient for both youth and parent raters. Parent reports on the 35-item, 12-item, and domain scores for the WHODAS 2.0 were lower than youth reports. There were significant differences between parent and youth scores for both versions of the WHODAS 2.0 scores, and for the cognition, mobility, and participation domains. Bland-Altman plots revealed measurement error between informants, and agreement was low to moderate (ICC -0.04 to 0.33). Logistic regression analyses revealed that household income <$75,000 was associated with lower odds of meaningful disagreement between parent and youth WHODAS 2.0 scores on the 36-item WHODAS 2.0, and increased youth age was associated with lower odds of meaningful disagreement between parent and youth WHODAS 2.0 scores on the 12-item WHODAS 2.0.

Conclusion: Because conclusions derived from both versions of the WHODAS 2.0 are similar, the 12-item version is sufficient for measuring functional impairment in a clinical context of youth with mental disorder. However, reports from both youth and parents appear valuable in understanding functional impairment. Additional work is needed to understand the factors that influence discrepancies and the implications for care.

Acknowledgements

First, I would like to thank my supervisor Dr. Mark Ferro. I am grateful to have had the opportunity to learn from you through this degree. The efforts you make to support all of your students, especially the efforts you made to support us amidst COVID-19, are both noticed and appreciated. I have become a better researcher because of you.

Thank you to my committee members Dr. Chris Perlman and Dr. Sharon Kirkpatrick. My thesis is stronger as a result of your contributions. I am thankful I was able to learn from you in my graduate studies and I look up to you both as mentors.

I would also like to thank my family – Mom, Dad, Natasha, and Aiden. Thank you for teaching me the importance of hard work and simultaneously reminding me that life is about so much more. I am so thankful to have your support and love in my life!

To my epi girls – Nicole, Megan, Alex, and Pam. I am so blessed to have such strong, kind, and smart women as friends. Our weekly conversations are the truest demonstration that friendship is about much more than proximity (though I am excited for a time when we can all go out together again). Thank you all for filling this degree with laughter, encouragement, and friendships I know will last far beyond our studies.

And finally, to my fiancé, work from home partner, best friend, and life buddy Saad. Thank you for loving and encouraging me in all that I do. I am so thankful that you push me to be the best version of myself and chase my dreams. There is no one else I would rather share my life with. I am excited for all that the future holds for us!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

World Health Organization Disability Assessment Schedule	WHODAS
Attribution Behaviour Context	ABC
Mini International Neuropsychiatric Interview for Children and Adolescents	MINI-KID
Perceived Stress Scale	PeSS
Center for Epidemiologic Studies Depression Scale	CES-D
State-Trait Anxiety Inventory	STAI
Intraclass correlation coefficient	ICC

BACKGROUND

1.1. Mental disorders in youth

Worldwide approximately one in five youth have been affected by a mental disorder within the previous year (1). This finding corroborates evidence from a 2015 meta-analysis exploring the worldwide prevalence of mental disorder, which found over 241 million (14%) children and youth were currently experiencing mental disorder (2). Within Ontario, it is estimated 22% of youth report experiencing a mental disorder within the previous 6-months (3). Youth are an important cohort to consider, as the prevalence of mental disorder in Canada is highest among this group (4). While a 2019 investigation concluded that the prevalence of mental disorders in youth was stable between 1983 and 2014, it highlighted that there is an increased perceived need for help and utilization of mental health services (5,6). For example, in 2012, 12% of youth aged 15-24 sought professional mental health support (4).

The short-term impact of these disorders is profound as mental disorders place a considerable burden on youth themselves, their family, and the healthcare system overall. While the signs and symptoms of mental disorders vary depending on the specific diagnosis, mental disorders can be troubling for youth, resulting in both functional impairment and decreased health-related quality of life (7). Additionally, mental disorders in youth are three times more likely to have a severe impact on school, family, and social functioning compared to the impact of mental disorders in children (8). Considering the impact of youth mental disorder more broadly, within the United States, treatment for mental health and substance use concerns represents 9% ($9.6 billion USD) of healthcare expenditure on children and youth (9). Within Canada mental illness costs the economy approximately $50 billion each year, and within Ontario indirect costs associated with youth mental health are over $421 million annually (10,11).

In addition to short-term considerations, the impacts of mental disorder also extend across the life course. First, mental disorder can be chronic in nature; the majority of mental disorders in adulthood will emerge by or within adolescence (12). Mental disorders in youth are associated with lower educational achievement, substance abuse, violence, reproductive, and sexual health concerns (13). Additionally, compared to those without mental disorders, individuals who reported a mental disorder as youth had increased problematic social relationships, poorer psychological

health, increased adversity in their environmental context, and decreased quality of life 17 years later (14).

As the functional impairment of mental disorder increases, the frequency and intensity of symptoms, as well as the limitations associated with symptoms also increases. Functional impairment moderates many of the negative effects of mental disorders (15–17). For example, individuals with substantial functional impairment as a result of mental disorder are more likely to have decreased quality of life and increased healthcare utilization, compared to those with less functional impairment and those who are healthy (15). Thus, to minimize the impact of mental disorders, monitoring of functional impairment is a clinical priority often accomplished using health assessments.

1.2. Health Assessments

The assessment of mental disorders is complicated by their invisibility; valid and reliable physiologic measures are often not available to monitor a mental disorder. Additionally, functional impairment is considered a latent construct as it cannot be directly observed. To overcome this challenge, some investigations consider objective surrogate measures of functional impairment, such as health expenditure (18). While surrogate measures of functional impairment may provide some information, patient-reported scales specifically designed to assess functional impairment are integral for clinicians to evaluate the health and functioning of youth with mental disorder (19).

Health assessments often require respondents to subjectively quantify a characteristic or clinical judgement (20). Scale items may ask respondents to assign a numeric score or choose one response from an ordinal group. Multi-item scales ask various questions about observable elements that are related to the overall construct being assessed (21). It has been demonstrated that responses to these scales are valid, reliable, and may provide more accurate assessments of the underlying construct compared to single-item scales (20,22). Once generated, the psychometric properties of health assessments must be considered to ensure that the measurement tool is both valid and reliable within the population of interest.

The value of a scale differs depending on the context in which it is administered. Functional impairment assessments are valuable to clinicians, as they help monitor patient progression and elucidate if treatment is effective (23,24). However, within the time and resource constraints of most clinical settings, clinicians may choose not to administer a scale if it is time-consuming or

lacks versatility (25). In addition to the importance of scale versatility within a clinical setting, scales that can be administered to diverse groups are important within research settings. Scale versatility makes it possible for between-group comparisons where respondents, with varied health conditions or ethnicities, can be compared. Additionally, while participant fatigue should be considered, research settings may allow more time for respondents to complete assessments deemed 'too long' to have clinical utility.

1.3. WHODAS 2.0

The World Health Organization Disability Assessment Schedule (WHODAS) 2.0 is a 36-item questionnaire that comprehensively assesses functional impairment by considering six domains of functioning: cognition, mobility, self-care, getting along, life activities, and participation (26,27). A shorter 12-item version of the WHODAS 2.0 is also available and includes items that explain 81% of the variance of the 36-item version (28). Initially developed for use in psychiatry among adult populations, the WHODAS 2.0 is versatile as it is applicable to be used across physical and mental disorders (26). Most research considering the WHODAS 2.0 is focused on the 36-item version (29).

The psychometric properties of the WHODAS 2.0 have been evaluated using samples from different populations with varied health conditions and cultures; these tests consistently demonstrate the WHODAS 2.0 as a valid and reliable measure in diverse samples (30). Kimber et al. suggested that the psychometric properties of the 12-item WHODAS 2.0 in youth populations as young as 15 are also robust (29). The 36-item WHODAS 2.0 was also found to have adequate psychometric properties, in a sample Chinese youth between 10 and 19 years of age (31). Previous work has also demonstrated that valid comparisons of overall disability, measured with the WHODAS 2.0, can be made with confidence across youth aged 15-19 years with differing types of chronic illnesses (28).

1.4. Informants

Undoubtedly, within pediatric care, parents are important stakeholders. Parents are responsible for medical decision-making, and are often the party that initiates medical appointments (32). As a result, historically, health assessment scales have been completed by parent proxies who report on their child's health status (33). Some considered it unnecessary to

ask for child self-reports, believing that the developing cognitive and linguistic abilities required to accurately complete health assessments limited the ability for children to provide valuable reports (33). A growing number of psychometric evaluations which conclude that scales can be confidently administered to children and youth, reduce these concerns (34–36). Further, considering a youth population, concerns about youth's ability to understand and confidently complete questionnaires may be minimal (37). Another factor that may reduce the ability to collect self-reported information exists as children may be too unwell or unwilling to participate in the completion of health assessments, making parent report necessary (37). However, family-centered approaches to care place considerable value on self-reports from children and youth, and advocate that whenever possible these reports are used to inform decision-making and the evaluation of outcomes (38).

Contradictory to the assumption that parent proxy-report accurately summarizes child and youth mental health, initial studies considering both parent and child informants revealed poor agreement (39). Discrepant reports are complex to manage as a single 'correct' respondent is not available for comparison (40). To date, much work on informant discrepancies has focused on explaining factors associated with disagreement. It is well established that disagreement is reduced when reporting observable characteristics or behaviours (39,41). In the context of mental disorder, agreement may be higher on externalizing disorders compared to internalizing disorders. While some studies support this (39,42–44), others conclude that disagreement does not differ between internalizing and externalizing disorders (45).

In addition to the disorder itself, youth characteristics such as child sex and comorbid health concerns may influence agreement. Some investigations report that child sex influences agreement (34,46); however, others conclude that this is not significantly associated with decreased agreement (32). Additionally, the presence of a comorbid physical health condition increases parent-child discrepancies (46,47). Parent characteristics are also known to influence disagreement; increased maternal stress (42,48), and parent psychopathology are independently associated with disagreement (46).

Considering a youth population specifically, discrepancies may be more meaningfully explained by additional factors. First, youth have expanded social networks. Youth who confide in peers may be less likely to regularly inform parents about mental health concerns. Thus, parents may not be fully aware of the impact or functional impairment of their child's mental disorder

(46). Additionally, as children become older parents encourage youth to become responsible for their health, and are less likely to closely monitor health conditions that are not life threatening (23,49). Youth report that they are less likely to expect parental involvement and are more likely to delay disclosing information until support is required (23). Overall, parents may have less information about the mental health status of youth resulting in discrepant reports.

The Attribution Behaviour Context (ABC) model provides a useful theoretical framework to explain and understand informant discrepancies (40). The ABC model draws upon psychological phenomena such as actor-observer bias, attribution theory, recall and source monitoring to explain why respondents may provide discrepant reports (40). This model suggests that each informant has a unique perspective that is influenced by the respondent's individual characteristics (40,46). These unique perspectives influence what individuals attribute as the causes of behavior, which alter their responses on health assessments, ultimately leading to discrepancies (40).

The ABC model is useful in explaining why reports from multiple informants are valuable. It is widely acknowledged that each informant provides information that can be meaningful (32,47). While some groups advocate that youth, report is essential and parent report should be used to compliment findings (37), many argue that collecting information from both informants is ideal (32,47,50).

Unfortunately, discrepancies between parent and youth reports can have profound impacts on the effectiveness of care and the supports provided to youth with mental health disorders (51). For instance, parental underestimates of youth disease severity have been shown to reduce the amount of supportive care provided to youth, despite the need for such services (19,52). Additionally, high disagreement between parent-youth dyads is associated with poorer health outcomes for youth (53). Without understanding the extent to which informant discrepancies (parent vs. youth) exist, care plans and treatment decisions for youth with mental disorder may be negatively impacted (33,49,54).

STUDY RATIONALE AND RESEARCH OBJECTIVES

Research examining informant discrepancies – particularly differences between parent and youth reports – is lacking. Evidence shows that agreement between parent and child reports of child mental health and quality of life is low, but the extent to which this phenomenon extrapolates to other aspects of children's health (i.e., functioning and disability), is unknown (49,55). Additionally, despite the value multiple informant responses may provide, collecting this information may not be feasible in clinical settings. Therefore, identifying the factors most associated with youth-parent discrepancies will be valuable and can be taken into account by clinicians.

The specific aims of this thesis were to:

1. Estimate parent-youth agreement on the WHODAS 2.0 in a clinical sample of children with mental disorder(s).

It was hypothesized that parent-youth agreement would be low to moderate, and that parents would report lower WHODAS 2.0 scores compared to youth. Agreement was expected to be similar between the 12-item and 36-item WHODAS 2.0 versions. It was also expected that the *mobility* and *self-care* domains would have higher agreement that other domains, as these domains explore more observable characteristics than the other WHODAS 2.0 domains.

2. Examine convergent and divergent validity, as well as internal consistency reliability of the WHODAS 2.0 between informants.

It was hypothesized that KIDSCREEN domains would be strongly correlated with the 12-item and 36-item WHODAS 2.0 scores. Conversely, it was not anticipated that youth age, youth sex, and household income would be strongly correlated with WHODAS 2.0 scores. It was expected that the strength of correlations would be similar between parent and youth reports. Additionally, it was predicted that internal consistency would be sufficient for both parent and youth informants.

3. Examine factors for clinically meaningful disagreement between parent vs. youth report.

It was hypothesized that demographic, youth health, parent health, and/or healthcare setting variables may be useful at explaining parent-youth disagreement on the WHODAS 2.0. It was also expected that the covariates that best explain the variation between parent-child disagreement would be similar for the 36-item and 12-item WHODAS 2.0.

METHODS

3.1. Study Design and Sample

The data for this thesis comes from a cross-sectional study that investigated the prevalence of multimorbidity (co-occurring mental and physical illness) among children and youth aged 4-17 years, and explored factors associated with multimorbidity and health service use. Participants were recruited from a pediatric academic tertiary care teaching hospital in Ontario.

Overall, 259 eligible youth were identified, 144 provided informed consent, and 92 dyads completed the study. Only participants (herein youth) aged 14–17 years completed the WHODAS 2.0 at the in-person interview; analyses in the current investigation were completed using data from a subset of participants. Overall, 66 (72%) parent-youth dyads (youth between 14 and 17 years of age) were included within this work.

Youth were eligible to participate in this study if they: 1) were between 14 and 17 years of age; 2) were classified as having major depressive disorder, separation anxiety, phobia, generalized anxiety, attention-deficit/hyperactivity disorder, conduct disorder or oppositional defiant disorder; 3) were currently receiving inpatient or outpatient mental health care; 4) had a parent who was their primary caregiver for the three months prior to referral for mental health care, and; 5) both youth and parent had sufficient command of the English language to complete the study questionnaires. The maximum age of participants was selected as it is the maximum age of youth eligible to receive health care services at a pediatric institution within Canada. Further, the disorders of interest were selected as research shows these are the most common youth mental disorders (1,8). The necessity for a consistent primary caregiver was selected as the study aimed to compare parent-report to youth-report, this would not be possible if the primary caregiver role was recently assumed. English proficiency was required as most measures used in the study have not be validated in other languages. As they would not be fully able to participate, youth diagnosed with schizophrenia or those exhibiting symptoms of psychosis were not eligible to participate in the study.

3.2. Procedure

This study received all relevant ethical approvals. Research staff confirmed that potential participants met the age-eligibility and English language proficiency required; they provided families with a study information letter which included information about the studies objectives

and participation requirements. Families who maintained interest were then contacted by the Study Coordinator who administered The Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID) to determine if a youth met the criteria for having major depressive disorder, separation anxiety, phobia, generalized anxiety, attention-deficit/hyperactivity disorder, conduct disorder, or oppositional defiant disorder.

An in-person interview for parents and youth was scheduled for those who screened positive for at least one mental disorder. At this visit, the Research Coordinator obtained written informed consent for participation in the study and data were collected through computer-assisted self-reports separately from youth and parents. This visit was approximately 45 minutes for youth and 60 minutes for parents. Additional information about the study is available elsewhere (56).

3.3. Sample Size Considerations

Overall, there is no consensus about how sample size for studies examining psychometric properties of scales should be calculated a priori (57,58). However, the sample size needed to compare reliability ratings of continuous measures (such as WHODAS 2.0 scores) are smaller than those required to compare binary ratings (59). The sample size utilized within this study is similar to other investigations of parent-youth agreement (n= 62) (47), and is comparable to the size requirements that previous reports suggest is sufficient, n= 30–40 (60), n= 60/group (59). Additionally, a literature review exploring patient reported outcomes found that 92% of studies utilized a subject:item ratio, a measure often used to justify sample size, ≥ 2 (58). Within this study the subject:item ratio is 1.83 and 5.5 for the 36-item and 12-item versions of the WHODAS 2.0, respectively (58). Therefore, the sample size utilized in this study is likely to be widely regarded as sufficient.

3.4. Measures

Sociodemographic variables (such as youth age, youth sex, parent age, household income) and clinically relevant variables (such as care setting, comorbid physical conditions) were collected.

3.4.1. MINI-KID

The MINI-KID is a short, structured, diagnostic interview that utilizes skip patterns and screening questions to identify the presence of 24 child and adolescent mental disorders from the DSM-IV and ICD-10 (61). Respondents are instructed to consider the previous six months when answering questions. Within each module, the main diagnostic criteria for the respective mental disorder are assessed. If these screener criteria are not met the module is skipped, and the next mental disorder is assessed. If screener criteria are met, respondents are asked all questions within the module to assess if the mental disorder is present. The time to complete the MINI-KID varies depending on the number of modules that are completed.

While the questions asked to parents and youth were the same, the subject of the questions varied. For example, youth were asked questions where the subject was denoted as 'you' while the parental version specified 'he/she'.

The MINI-KID has strong psychometric properties, and is regarded as valid and reliable in clinical and general populations (62). Using the *Schedule for Affective Disorders and Schizophrenia for School Aged Children-Present and Lifetime Version*, the MINI-KID has been validated; factor analysis looking at similar and disparate constructs suggests that convergent and divergent validity, between the MINI-KID and the *Brief Child and Family Phone Interview*, is present (62). The parent proxy of the MINI-KID has good diagnostic agreement (κ = 0.46-0.94), interrater reliability (AUC $\geq$ 0.97), and test-retest reliability (AUC $\geq$ 0.87) among outpatients (61). Additionally, the test-retest reliability of the MINI-KID is acceptable for both youth (κ = 0.57) and parent proxy-report (κ = 0.66) (62).

3.4.2. WHODAS 2.0

There are two versions of the WHODAS 2.0 (26). The long version of the WHODAS 2.0 is 36-items and takes 5-10 minutes to complete (26). The WHODAS 2.0 measures 6 domains: cognition (6-items), mobility (5-items), self-care (4-items), getting along (5-items), life activities (8-items) and participation (8-items). A 12-item version of the WHODAS 2.0 is also available and considers the items that most strongly loaded onto the domain-factor structure in the 36-item version (63). For these analyses, both the 36-item and the 12-item scores will be considered. Respondents answer questions by reflecting on the amount of difficultly they (or their child) had

within the previous 30 days. Questions are answered on a 5-point scale (1- none, 2- mild, 3- moderate, 4- severe, and 5- extreme of cannot do) or respondents can select 'do not know' or 'not applicable' responses. Similar to the variation between parent and youth versions of the MINI-KID, the questions asked on both versions of the WHODAS 2.0 varied only by subject. The youth version included questions where the subject was denoted as 'you' while the parent version asked questions where the subject was 'your child'.

Two scoring methods – simple and complex – can be utilized to determine WHODAS 2.0 scores. The simple method can be applied to generate overall scores on both the 12-item and 36-item versions and domain scores. Using this approach, scores are determined by taking the sum of all answers provided and then dividing by the total number of items considered. Therefore, the simple scoring method generates scores between 1 and 5. The complex method of scoring utilizes item-response-theory based scoring. Within the complex scoring method, the response for each item is considered, a summary score for each domain is generated through differentially weighting items and the level of severity specified. This scoring method generates overall scores and domain-specific WHODAS 2.0 scores. Scores range from 0 (denoting no disability) to 100 (denoting full disability). Regardless of the scoring method applied or the score considered (domain -specific or overall), higher scores indicate increased disability. Within this analysis the simple scoring method will be used, as this method would be employed to generate scores in a clinical (compared to research) setting. This scoring method is valuable to ensure the results of this work provide meaningful information to health professionals. The simple method of scoring will be used to compute several WHODAS 2.0 scores: the 12-item WHODAS 2.0 overall score, 36-item WHODAS 2.0 overall score, and 36-item WHODAS 2.0 domain-specific scores.

In populations of individuals with mental disorder(s), Cronbach's α for overall WHODAS 2.0 score was 0.98 (domain-specific scores Cronbach's $\alpha = 0.92 – 0.94$) (30). This finding was supported by a review which identified the 12-item WHODAS 2.0 as reliable, internally consistent, and strongly correlated with other measures of disability (Cronbach's $\alpha = 0.81 – 0.96$) (64).

3.4.3 KIDSCREEN

The KIDSCREEN is a 27-item questionnaire exploring child and youth health-related quality of life on five dimensions: physical well-being (5-items), psychosocial well-being (7-items), autonomy and parents (7-items), peers and social support (4-items), and school

environment (4-items). The questionnaire takes 5-10 minutes to complete, and questions are asked on a 5-point Likert scale (0- never, 1- seldom, 2- quite often, 3- very often, and 4- always). Four items are reverse coded. Higher scores indicate greater health-related quality of life. Raw scores are transformed into T-values with a mean of 50 and a standard deviation of 10. Parent and youth versions of the KIDSCREEN include identical questions. The parent version includes questions with the subject 'your child' while the youth version utilizes 'you' as the subject.

The KIDSCREEN has a replicable factor structure across countries (65). Recent findings among youth with mental disorder(s), show that the KIDSCREEN is a valid measure of health-related quality of life and demonstrates partial measurement invariance between parent-youth dyads (66). Construct validity and criterion validity between the KIDSCREEN and other health-related quality of life measures have also been established (67).

3.4.4. Perceived Stress Scale (PeSS)

The PeSS is a 10-item questionnaire assessing the perception of stress within youth which takes less than 5-minutes to complete (68,69). Questions ask about thoughts and feeling within the previous month and answers are provided on a 5-point scale (0- never, 1- almost never, 2- sometimes, 3- fairly often, and 4- very often). Raw scores are summed to determine a final score. Higher scores indicate an increase in life situations being perceived as stressful. The PeSS has adequate reliability, has been correlated with other assessments of youth stress, and has convergent and divergent validity with scales assessing similar and disparate constructs, respectively (69,70). The factor structure, internal consistency, and convergent validity of the PeSS have been explored in a sample of parents who have children with mental illness; the PeSS was found to be valid and reliable in this sample (71).

3.4.5. Center for Epidemiologic Studies Depression Scale (CES-D)

The CES-D is a 20-item measure of depressive symptoms within the previous week that takes approximately five minutes to complete (72). Parents within this study are asked to report how often they felt a particular symptom on a 4-point scale (0- rarely, 1- sometimes, 2- occasionally, 3- most of the time). After reverse coding 4 questions, items are summed. Higher scores indicate greater symptoms of depression. While the factor structure of the CES-D differs across racial and ethnic groups (73), it is recognized as a reliable and valid measure of depressive

symptoms (74). The psychometric properties of CES-D are consistent across various adult populations (75), including a sample of parents who have children with mental illness (71).

3.4.6. State-Trait Anxiety Inventory (STAI)

The STAI is a 40-item measure asking about general feelings of anxiety (76). In this study, only the 20 items focusing on trait anxiety (how individuals generally feel), were completed by parents. The STAI takes 10-minutes to complete, and respondents are asked to agree or disagree with a statement on a 4-point scale (1- not at all, 2- somewhat, 3- moderately so, and 4- very much so). Final scores are computed by reverse coding 7 questions and then taking the total sum of all answers. Higher scores indicate greater anxiety. The psychometric properties of the STAI have been extensively tested in diverse populations (77), including a sample of parents who have children with mental illness (71). Test-retest reliability and internal consistency have been routinely demonstrated (76,78).

3.5. Analysis Plan:

All analyses were performed in SAS Studio, Version 9.0.4 (SAS Institute, Cary, NC). Hypothesis tests were two-sided with $\alpha = 0.05$ applied.

3.5.1 Missing Data

It was possible that parent and youth responses on the WHODAS 2.0 may not be complete. In the case where responses were missing, imputation methods were employed. Specifically, the number of missing responses was considered for each of the WHODAS 2.0 domains. If more than 50% of items in any domain were missing, the dyad was removed from all analyses. However, if at least 50% of the items had been answered, the average domain score was computed and assigned to those items missing a response.

3.5.2 Objective 1: Estimate parent-youth agreement on the WHODAS 2.0 in a clinical sample of children with mental disorder(s).

Descriptive statistics for overall and domain-specific WHODAS 2.0 scores were summarized separately for parent and youth groups. It was not expected that parent and youth

reports would be equivalent; however, as these informants are providing information on the same subject (i.e. youth disability), a high correlation coefficient between parent- and youth-report was expected. Regardless, the strength of an association does not provide information about the agreement between two measures (79). To adequately assess agreement between parent-report and youth-report on the WHODAS 2.0, the limits of agreement, and intraclass correlation coefficient (ICC) were computed, and the Bland-Altman method was employed. A paired t-test was also completed to discern if there were statistically significant differences between the mean of youth-reported WHODAS 2.0 scores and the mean of parent-reported WHODAS 2.0 scores. These analyses were repeated for overall WHODAS 2.0 scores and domain-specific scores.

Bland-Altman plots were constructed by plotting the difference between parent-report and youth-report on the Y-axis. It is unclear what the 'true' WHODAS 2.0 score for youth is; therefore, the average score between informants was calculated and plotted on the X-axis. The line of identity, a horizontal line at Y= 0 which indicates where points would fall if there was complete agreement between parent- and youth-report, was also plotted (79). Each parent-youth dyad represented one point on the Bland-Altman plot. Systematic differences between parent- and youth-report were assessed by visually inspecting the Bland-Altman plot. If the points were equally distributed above and below the line of identity and no discernable pattern was present (i.e. the average of the differences was close to zero), it was concluded that there was no systematic difference between parent-report and youth-report (79). If there were deviations, such that points were not equally and randomly scattered around the line of identity (i.e. the average of the differences was not close to zero), it was concluded that there was a systematic difference between parent- and youth-reports (79).

The limits of agreement were also calculated, using the average and standard deviation of the differences, and added to the Bland-Altman plots. Limits of agreement include both systematic and random error and allow the differences between parent- and youth-reports to be compared (79,80). If the limits of agreement were narrow, the measurement error between parent-report and youth-report was considered small. However, if the limits of agreement were wide, the measurement error between parent-report and youth-report was large. While narrow limits of agreement were favorable, as they indicate parent-report and youth-report have high agreement, it is important to note that there is no statistical test to indicate what amount of disagreement is clinically meaningful (79).

To determine the limits of agreement, a horizontal line denoting the average of the differences was plotted. The upper and lower limits of agreement were determined by adding and subtracting, respectively, 1.96 standard deviations of this value to the average of the differences (79,80). By definition, 95% of all points fell between the limits of agreement (79).

In addition, the ICC was computed for both parent and youth, overall and domain specific, WHODAS 2.0 scores. The ICC is a measure of reliability between two different informants evaluating one construct. High inter-rater reliability, demonstrated with a high ICC, indicates the scale is not subject to measurement error due to changes in raters. With respect to this thesis, ICC measurements indicated the degree of reliability between parent and youth reports of youth disability. As identified by Shrout and Fleiss, a 2-way mixed-effect model for ICC was appropriate to assess interrater reliability within this analysis as were multiple questions answered by a single rater (81). Correlations were evaluated as suggested by Koo et al., ICC < 0.50 as poor reliability, ICC = 0.50 – 0.75 as moderate reliability, ICC = 0.75 – 0.90 as good reliability and ICC > 0.90 as excellent reliability (82).

3.5.3. Objective 2: Examine convergent and divergent validity, as well as internal consistency reliability of the WHODAS 2.0 between informants.

Convergent validity was assessed by considering the correlation between WHODAS 2.0 scores and each of the KIDSCREEN domains – physical well-being, psychological well-being, autonomy and parents, peers and social support, and school environment. Divergent validity was assessed by considering the correlation between WHODAS 2.0 scores and demographic factors – youth age, youth sex, and household income. In all comparisons, youth-reports of the WHODAS 2.0 were compared to youth-reports of the KIDSCREEN and parent-reports of the WHODAS 2.0 were compared to parent-reports of the KIDSCREEN.

Convergent validity was evaluated by separately considering the correlations between the KIDSCREEN domains with overall WHODAS 2.0 scores (both 36-item and 12-item). As these domains represent similar latent constructs, a strong correlation would suggest convergent validity. Divergent validity (or discriminant validity) was assessed by separately computing the correlation between youth age, youth sex, and household income. These variables were not expected to be related to functional impairment (as measured by the WHODAS 2.0). Thus, the absence of strong correlations would suggest divergent validity.

Kendall's Tau correlation coefficient (τ) was used to quantify the correlation between WHODAS 2.0 scores and continuous variables (KIDSCREEN domains and youth age). τ is a conservative alternative to Spearman's correlation coefficient, and appropriate to assess concordance or discordance between pairs of observations (83). Evaluation of correlations were assessed such that, $\tau = |0.1 – 0.29|$ was considered a weak correlation, $\tau = |0.3 – 0.49|$ as a moderate correlation, and $\tau \geq |0.5|$ as a strong correlation (84). A point biserial correlation coefficient was calculated to quantify the correlation between WHODAS 2.0 scores and binary variables (youth sex and household income). Point biserial correlation coefficients were interpreted similarly to τ correlation coefficients. Analyses were completed separately for parent- and youth-reports. Correlations were compared to assess if the degree of convergent and divergent validity differs between parent- and youth-reports.

Internal consistency was examined by computing ordinal α, a widely used reliability coefficient that appropriately accounts for data collected from ordinal scales such as the WHODAS 2.0 (85). Ordinal α was calculated for parent and youth reports for overall WHODAS 2.0 scores. An ordinal $\alpha > 0.70$ indicated that internal consistency was sufficient (86).

3.5.4. Objective 3: Examine factors for clinically meaningful disagreement between parent vs. youth report.

The point at which informant discrepancies become clinically meaningful cannot be derived simply through statistical analyses (79). A clinically meaningful important difference score for the WHODAS 2.0 has not yet been determined (26). To determine which parent-youth dyads had clinically meaningful differences between their WHODAS 2.0 scores, the distribution of the differences in parent-reported and youth-reported overall WHODAS 2.0 scores was considered. Dyads that had differences greater than 0.5 standard deviations from the mean were considered to have clinically meaningful differences in overall WHODAS 2.0 scores. This approach was consistent with research that explored agreement in health related quality of life (87).

Several covariates may increase the odds that a parent-youth dyad has clinically meaningful differences. These covariates include: 1. demographic information – such as youth age, youth sex, and family income, 2. youth health information – such as multimorbidity, and

perceived stress (*PeSS* score), 3. parental psychosocial distress (*STAI* and *CES-D* scores), and 4. care setting.

Binary logistic regression was used to model the association between covariates and clinically meaningful differences in overall WHODAS 2.0 scores between parent-youth dyads. Covariates of interest were the continuous and categorical exposure variables, and the presence of a clinically meaningful difference in overall WHODAS 2.0 scores was the dichotomous outcome variable. Model generation was completed for overall scores on both the 12-item and 36-item WHODAS 2.0. The PROC LOGISTIC procedure was used to compute odds ratios and 95% confidence intervals. Four hypothesized main effects models were created, with each sequential hypothesized model adding an additional block of covariates (Appendix A). Demographic covariates were followed by the youth health block, the parent health block, and the health care block. The c-statistic was used to assess model fit. While there is no consensus about a specific threshold for acceptable c-statistic values, results above 0.5 indicated that the model was performing better than chance (88). The model that balanced parsimony while maximizing the c-statistic between 0.5 and 1, denoting the estimated area under the ROC curve, was classified as the optimal model.

RESULTS

4.1. Missing Data

When responses on the 36-item WHODAS 2.0 were considered, there was extensive missing data for one question that asked about difficulty in sexual activities. The sexual activities question was not answered by 80.3% of parents and 40.9% of youth. As a result, this item was removed from all analysis and a modified 35-item WHODAS was employed. The removal of the sexual activities question was completed in another assessment of the WHODAS 2.0 with youth populations (31). Additionally, the initial sample included 66 parent-youth dyads. After assessing missingness, 10 dyads were removed as there was more than 50% of responses on at least one WHODAS 2.0 domain. Therefore, the final analytic sample included 56 parent-youth dyads.

4.2. Sample Characteristics

The mean age of youth in the sample was 15.75 years (SD = 1.03), and the majority of youth were female (76.79%) and Caucasian (90.57%). The most common mental disorder among the sample of youth (n= 66) was major depressive disorder (80.36%), followed by generalized anxiety disorder (76.79%), social phobia (60.71%), oppositional defiant disorder (35.71%), separation anxiety (30.36%), attention-deficit/hyperactivity disorder (26.79%), specific phobia (17.86%), and conduct disorder (17.86%). There were approximately equal numbers of youth recruited from inpatient settings (55.36%) compared to outpatient settings.

The mean age of parents was 47.46 years (SD = 6.13), and the majority of parent informants were female (83.93%). Household income was equally represented, almost half of the sample had an annual household income <$75000 (46.43%). Additional details of the study sample can be found in Table 1.

Table 1: Characteristics of the study sample

	Mean (SD)
Youth age, years	15.75 (1.03)
Youth stress, PeSS	26.77 (8.00)
Parent age, years	47.46 (6.13)
Parent psychosocial distress	66.25 (16.88)
	n (%)
Female youth	43 (76.79)
Multimorbidity	15(26.79)
Caucasian	48 (90.57)
Mental disorder:	
Major depressive disorder	45 (80.36)
Generalized anxiety	43 (76.79)
Separation anxiety	17 (30.36)
Social Phobia	34 (60.71)
Specific Phobia	10 (17.86)
Attention-deficit/hyperactivity disorder	15 (26.79)
Oppositional defiant disorder	20 (35.71)
Conduct disorder	10 (17.86)
Female parent	47 (83.93)
Household income <$75,000	26 (46.43)
Inpatient	31 (55.36)

4.3. Objective 1

WHODAS 2.0 Scores: Table 2 contains results of the WHODAS 2.0 scores for parents and youth. Parent reports on the 35-item, 12-item, and domain scores for the WHODAS 2.0 were lower than youth reports. There were significant differences between parent and youth scores for the 35-item and 12-item WHODAS 2.0 scores, and for the cognition, mobility, and participation domains.

Table 2: WHODAS 2.0 scores for youth and parents

	Parent Score Mean (SD)	Youth Score Mean (SD)	p-value
35- item	2.32 (0.63)	2.63 (0.76)	**0.02**
12-item	2.40 (0.70)	2.69 (0.76)	**0.04**
Cognition domain	2.57 (0.82)	2.93 (0.87)	**0.02**
Mobility domain	1.71 (0.66)	2.27 (0.98)	**<0.001**
Self-care domain	1.69 (0.77)	1.97 (0.93)	0.06
Getting along domain	2.46 (0.77)	2.65 (1.01)	0.23
Life activities domain	2.89 (0.94)	3.03 (0.98)	0.41
Participation domain	2.63 (0.75)	2.94 (0.98)	**0.05**

Figure 1 shows the Bland-Altman plot for the average 35-item WHODAS 2.0 scores. The points were scattered equally above and below the line of identity. While there was no clear pattern upon visual inspection, the points were clustered towards the centre of the plot and did not extend to the limits of the x-axis or y-axis. The average of the differences was -0.31. The upper and lower limits of agreement were 1.55 and -2.17, respectively. Therefore, the range of the limits of agreement is 3.72. Low agreement between parent and youth reports was present, as the limits of agreement were wide.

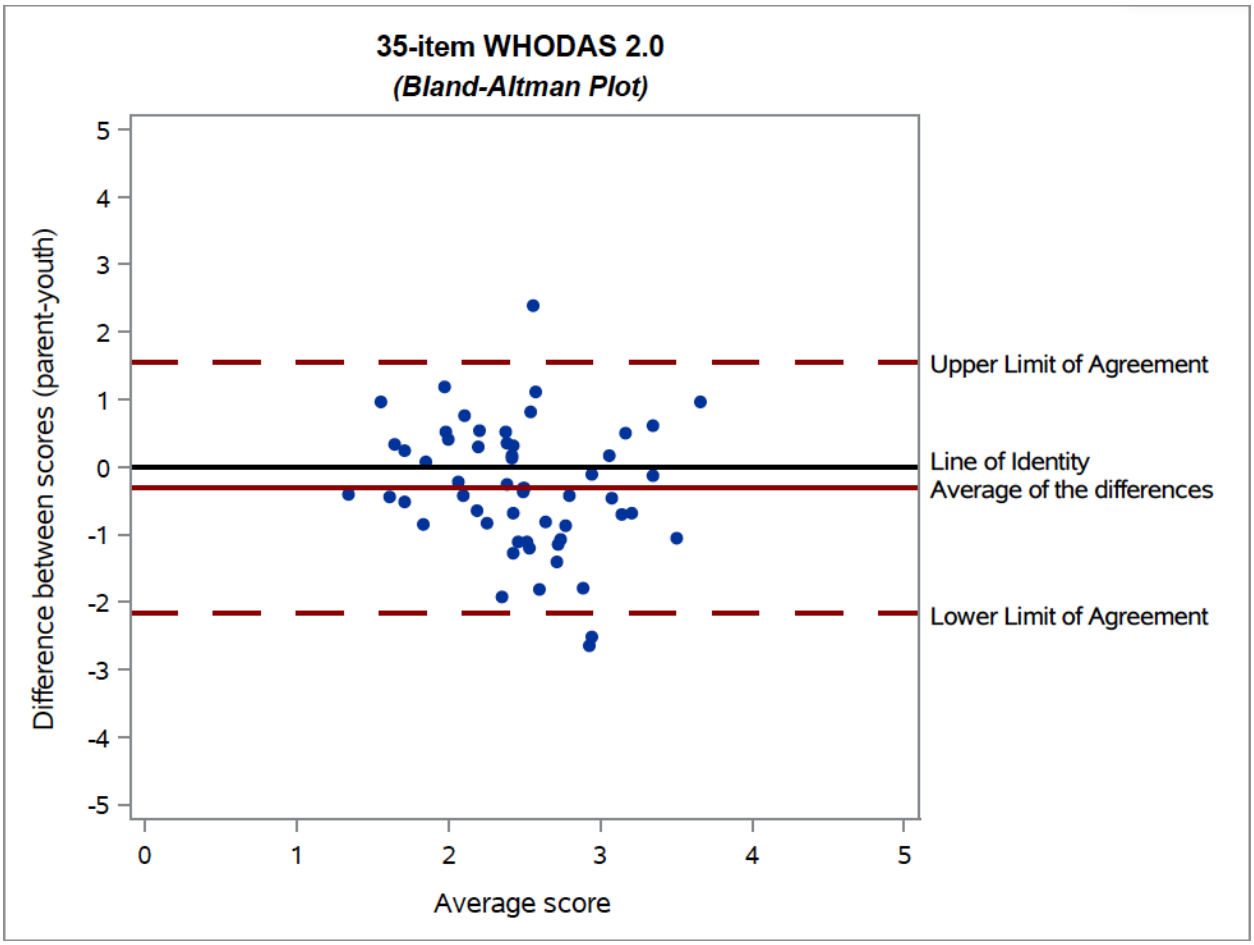


Figure 1. 35-item WHODAS 2.0 Bland-Altman Plot

Figure 2 shows the Bland-Altman plot for the average 12-item WHODAS 2.0 scores. The points were scattered equally above and below the line of identity. While there was no clear pattern upon visual inspection, the points were clustered towards the centre of the plot and did not extend to the limits of the x-axis or y-axis. The average of the differences was -0.29. The upper and lower limits of agreement were 1.69 and -2.27, respectively. Therefore, the range of the limits of agreement was 3.96. Low agreement between parent and youth reports was present, as the limits of agreement were wide.

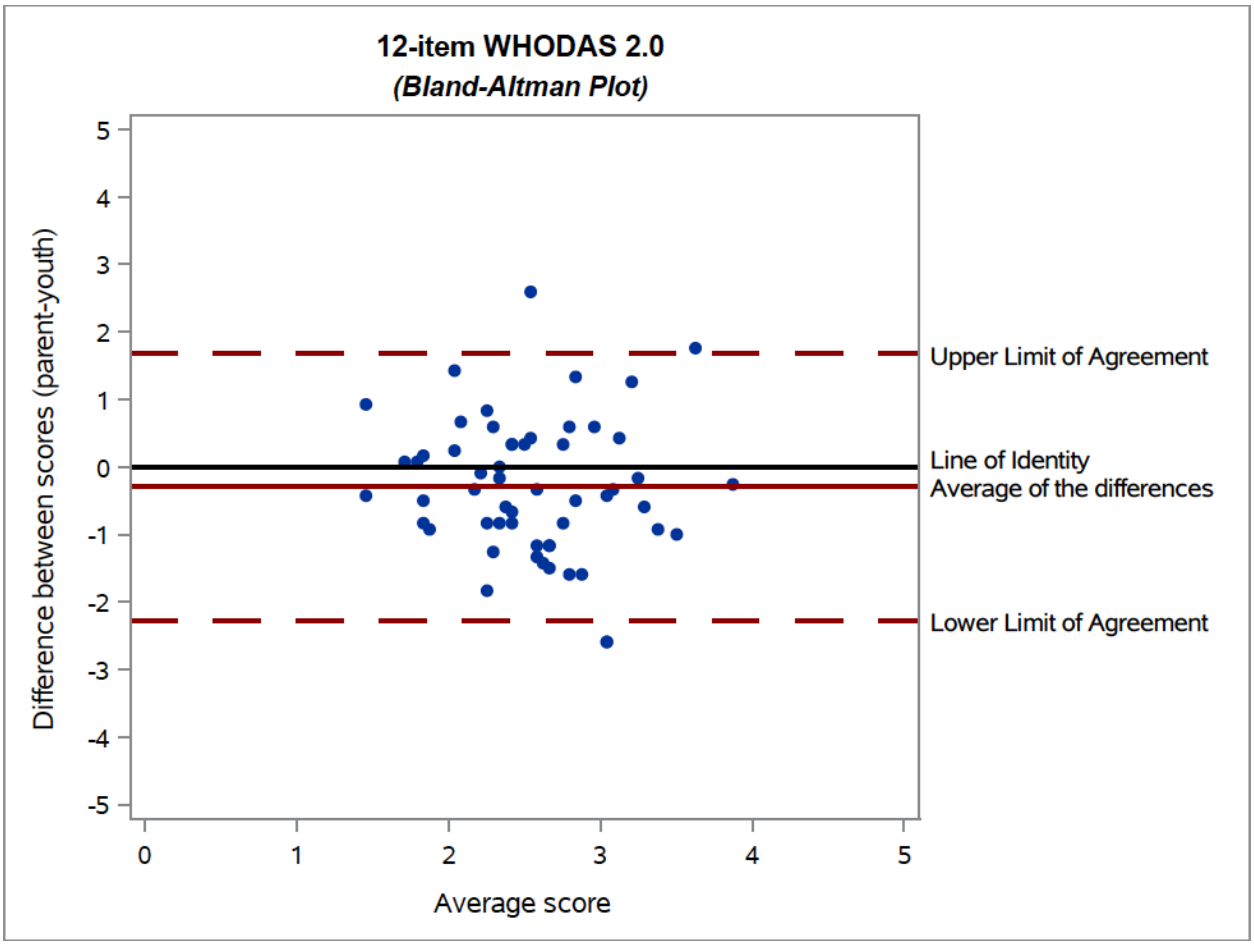


Figure 2. 12-item WHODAS 2.0 Bland-Altman Plot

The Bland-Altman plots for WHODAS 2.0 domain scores can be found in Appendix B. For all domains, points were scattered equally above and below the line of identity. There was no clear pattern for the cognition, getting along, life activities, and participation domains. For all domains the average of the differences were negative (-0.55 to -0.14). However, when considering the mobility and self-care domains variability increased across the x-axis. There may be increased variability between parent and youth responses on the mobility and self-care domains as the average WHODAS 2.0 score increased. Upon visual inspection, the range of the limits of agreement appeared wide across all domains (4.16 to 5.09). Thus, low agreement between parent and youth reports existed for all WHODAS 2.0 domains.

Table 3 contains the ICC results for the overall and domain WHODAS 2.0 scores. All ICC values were <0.5, indicating poor agreement between parent and youth reports. There was only one instance where the upper limit of the 95% CI was >0.5; the upper limit of the 95% CI for the self-care domain was 0.54, which would suggest that moderate reliability was possible. With the

exception of the self-care domain, the lower limit of the 95% CI was negative, as was the ICC point estimate for the cognition domain, indicating that the variability within a group of raters was greater than the variability between parent and youth raters.

Table 3: Parent-youth agreement

WHODAS 2.0 Score	ICC (95% CI)
35- item	0.07 (-0.19, 0.3)
12- item	0.02 (-0.24, 0.28)
Cognition domain	-0.04 (-0.29, 0.22)
Mobility domain	0.15 (-0.11, 0.40)
Self-care domain	0.33 (0.08, 0.54)
Getting along domain	0.13 (-0.14, 0.37)
Life activities domain	0.17 (-0.09, 0.41)
Participation domain	0.20 (-0.06, 0.43)

4.4. Objective 2

Table 4 contains the results of Kendall's τ correlation between the WHODAS 2.0 overall scores (on the 35-item and 12-item versions) and KIDSCREEN domains. When considering statistical significance ($p \leq 0.05$), conclusions were consistent between youth and parent reports for all correlations, with the exception of correlations including the *peers and social support* KIDSCREEN domain. There was not a significant correlation between the youth reported 35- or 12-item WHODAS 2.0 score and the *peers and social support* domain score (τ= -0.09 and -0.12, respectively). However, there was a significant correlation between the parent reported 35- or 12-item WHODAS 2.0 score and the *peers and social support* domain score (τ= -0.23 in both instances).

All correlations between WHODAS 2.0 overall scores and KIDSCREEN domain scores were low to moderate. Moderate correlation was seen when considering the youth reported *physical well-being* and *school environment* domains with both the 35- and 12-item youth reported WHODAS 2.0 scores. However, when considering parent reported scores, moderate correlations were only seen for the correlations that included the *school environment* KIDSCREEN domain and the correlation between the 12-item WHODAS 2.0 score and the *physical well-being* domain score. All other correlations for both parents and youth were low.

The direction of correlation was consistent for parents and youth with the exception of correlations considering the *psychological well-being* domain. A positive correlation was observed between youth reported *psychological well-being* and youth reported 35- and 12-item WHODAS scores (τ= 0.01 and <0.01, respectively). However, a negative correlation was observed between parent reported reported *psychological well-being* and parent reported 35- and 12-item WHODAS scores (τ= -0.07 and -0.05, respectively).

Table 4: Convergent validity

WHODAS 2.0 Version	KIDSCREEN Domain	Youth Report τ (p-value)	Parent Report τ (p-value)
35-item	Psychological well-being	0.01 (0.91)	-0.07 (0.45)
	Physical well-being	-0.31 (**<0.01**)	-0.24 (**0.01**)
	Parent relation and autonomy	-0.21 (**0.02**)	-0.29 (**<0.01**)
	Peers and social support	-0.09 (0.33)	-0.23 (**0.02**)
	School environment*	-0.41 (**<.0001**)	-0.40 (**<.0001**)
12-item	Psychological well-being	<0.01 (0.98)	-0.05 (0.65)
	Physical well-being	-0.31 (**<0.01**)	-0.30 (**<0.01**)
	Parent relation and autonomy	-0.21 (**0.03**)	-0.25 (**0.01**)
	Peers and social support	-0.12 (0.23)	-0.23 (**0.02**)
	School environment*	-0.41 (**<.0001**)	-0.42 (**<.0001**)

* n=55

Results for the correlation between youth and parent reported WHODAS 2.0 scores and demographic variables are reported in Table 5. When considering statistical significance ($p \leq 0.05$), conclusions were consistent between youth and parent reports for all correlations, with the exception of the correlation between the 12-item WHODAS 2.0 score and youth sex. There was a significant correlation between the youth reported 12-item WHODAS 2.0 score and youth sex (p-value= 0.03). However, there was not a significant correlation between the parent 12-item WHODAS 2.0 score and youth sex (p-value= 0.08).

All correlations between WHODAS 2.0 overall scores and demographic variables were low to moderate. Moderate correlation was seen between youth reported 35- and 12-item WHODAS 2.0 scores with youth sex, all other correlations were low.

The direction of correlation was consistent for parents and youth, with the exception of correlations considering youth age. A positive correlation was observed between youth reported 35- and 12-item WHODAS scores and youth age (0.03 and 0.07, respectively). However, a negative correlation was observed between parent reported 35- and 12-item WHODAS scores and youth age (-0.12 and -0.08, respectively).

Table 5: Divergent validity

WHODAS 2.0 Version	Demographic Variable	Youth Report Correlation* (p-value)	Parent Report Correlation* (p-value)
35-item	Female youth	0.32 (**0.02**)	0.29 (**0.03**)
	Youth age	0.03 (0.75)	-0.12 (0.22)
	Household income <$75,000	-0.03 (0.85)	-0.09 (0.51)
12-item	Female youth	0.30 (**0.03**)	0.23 (0.08)
	Youth age	0.07 (0.48)	-0.08 (0.42)
	Household income <$75,000	-0.06 (0.68)	-0.09 (0.49)

*****Kendall's τ for youth age correlations; Point biserial for female youth and household income <$75,000 correlations.

Internal consistency of the 35-item and 12-item WHODAS 2.0 was measured using ordinal α, for both parents and youth. The ordinal α for youth was 0.96 and 0.86, for the 35-item and 12-item WHODAS 2.0 versions, respectively. The ordinal α for parents was 0.96 and 0.91, for the 35-item and 12-item WHODAS 2.0 versions, respectively. As all ordinal α values were >0.7, the internal consistency of both versions of the WHODAS 2.0 was sufficient for both youth and parent raters.

4.5. Objective 3

The average of the differences between parent and youth reported 35-item WHODAS 2.0 scores was -0.31 (SD= 0.95). Therefore, dyads with a difference between -0.79 and 0.17 were not considered to have meaningful disagreement in the 35-item WHODAS 2.0 scores (n=19). Dyads outside that range were considered to have meaningful differences in the 35-item WHODAS 2.0 scores (n=37).

Table 6 shows the logistic regression results for models exploring factors associated with meaningful differences on the 35-item WHODAS 2.0. The only covariate that was significantly

associated with meaningful disagreement was annual household income <$75,000. Specifically, household income <$75,000 was associated with lower odds of meaningful differences between parent and youth 35-item WHODAS 2.0 scores, when controlling for: demographic factors (model 1, OR= 0.29, 95% CI= 0.09-0.97, c-statistic= 0.70); demographic factors and youth health covariates (model 2, OR= 0.29, 95% CI= 0.08-0.97, c-statistic= 0.71); and demographic factors, youth health, parent health, and health care setting (model 4, OR= 0.28, 95% CI= 0.08-0.99, c-statistic= 0.71). However, household income <$75,000 was not associated with lower odds of disagreement when controlling for demographic factors, youth health, and parent health (model 3, OR= 0.30, 95% CI= 0.09-1.04, c-statistic= 0.72). A crude model including only household income <$75,000 was also generated. This analysis revealed an OR= 0.36 (95% CI= 0.11-1.12) and a c-statistic= 0.63.

Table 6: Factors associated with meaningful differences on the 35-item WHODAS 2.0

Variables	Model 1 OR (95% CI)	Model 2 OR (95% CI)	Model 3 OR (95% CI)	Model 4 OR (95% CI)
Demographics				
Youth age	0.61	0.61	0.60	0.58
	(0.34, 1.12)	(0.33, 1.13)	(0.32, 1.10)	(0.31, 1.08)
Female youth	1.24	1.25	1.53	1.74
	(0.31, 4.90)	(0.31, 5.05)	(0.35, 6.69)	(0.38, 8.01)
Household Income	**0.29**	**0.29**	0.30	**0.28**
<$75,000	**(0.09, 0.97)**	**(0.08, 0.97)**	(0.09, 1.04)	**(0.08, 0.99)**
Youth Health				
Multimorbidity		0.88	0.90	0.74
		(0.23, 3.40)	(0.23, 3.51)	(0.18, 3.13)
Youth Stress		1.00	0.99	1.01
		(0.93, 1.08)	(0.92, 1.07)	(0.92, 1.10)
Parent Health				
Parent psychosocial distress			1.01	1.01
			(0.97, 1.04)	(0.98, 1.05)
Setting				
Inpatient				0.53
				(0.11, 2.56)
C-Statistic	0.70	0.71	0.72	0.71

The average of the differences between parent and youth reported 12-item WHODAS 2.0 scores was -0.29 (SD= 1.01). Therefore, dyads with a difference between -0.80 and 0.22 were not considered to have meaningful differences in the 12-item WHODAS 2.0 scores (n=18). Dyads outside that range were considered to have meaningful differences in the 35-item WHODAS 2.0 scores (n= 38).

Table 7 shows the logistic regression results for models exploring factors associated with meaningful differences on the 12-item WHODAS 2.0. The only covariate that was significantly associated with meaningful disagreement was youth age. Specifically, higher youth age was associated with lower odds of meaningful differences between parent and youth 12-item

WHODAS 2.0 scores, when controlling for: demographic factors (model 1, OR= 0.49, 95%CI= 0.26-0.93, c-statistic= 0.69); demographic factors and youth health covariates (model 2, OR= 0.46, 95% CI= 0.24-0.90, c-statistic= 0.72); demographic factors, youth health, and parent health covariates (model 3, OR= 0.44, 95% CI= 0.22-0.88, c-statistic= 0.73); and demographic factors, youth health, parent health, and health care setting (model 4, OR= 0.40, 95% CI= 0.19-0.84, c-statistic= 0.76). A crude model including only youth age was also generated. This analysis revealed an OR= 0.52 (95% CI= 0.28-0.98) and a c-statistic= 0.66.

Table 7: Factors associated with meaningful differences on the 12-item WHODAS 2.0

Variables	**Model 1** OR (95% CI)	**Model 2** OR (95% CI)	**Model 3** OR (95% CI)	**Model 4** OR (95% CI)
Demographics				
Youth age	**0.49**	**0.46**	**0.44**	**0.40**
	(0.26, 0.93)	**(0.24, 0.90)**	**(0.22, 0.88)**	**(0.19, 0.84)**
Female youth	1.36	1.25	1.39	1.69
	(0.34, 5.44)	(0.31, 5.05)	(0.32, 6.06)	(0.36, 8.03)
Household Income	0.45	0.41	0.46	0.38
<$75,000	(0.13, 1.52)	(0.12, 1.44)	(0.13, 1.64)	(0.10, 1.45)
Youth Health				
Multimorbidity		0.80	0.85	0.57
		(0.21, 3.15)	(0.21, 3.42)	(0.12, 2.66)
Youth Stress		1.04	1.04	1.07
		(0.96, 1.12)	(0.96, 1.13)	(0.97, 1.17)
Parent Health				
Parent psychosocial distress			0.98	0.99
			(0.95, 1.02)	(0.95, 1.03)
Setting				
Inpatient				0.33
				(0.06, 1.88)
C-Statistic	0.69	0.72	0.73	0.76

DISCUSSION

5.1. Objective 1

The hypothesis that parent-youth agreement would be low to moderate, with parents reporting lower WHODAS 2.0 scores compared to youth, was supported. The hypothesis that mobility and self-care domains would have higher agreement, compared to other WHODAS 2.0 domain scores, was partially supported. While the self-care domain had the highest agreement, agreement on the mobility domain was lower than the agreement observed in the life activities and participation domains. As hypothesized, agreement was similar for the 12-item and 35-item versions of the WHODAS 2.0.

Results from Bland-Altman plots and ICC values, suggest that parent-youth agreement is low to moderate. Agreement tends to decrease when providing reports about non-observable characteristics, such as the functional impact of mental disorder (89). These findings align with other studies exploring parent-youth agreement. Low to moderate agreement between parents and youth has been demonstrated when considering both the presence of mental disorder (32,45,90,91), and the impact of mental disorder (37,41,87). Not only do these results align with literature, they support the ABC model's assertion that parents and youth each provide valuable, but often discrepant, perspectives (32,37,42).

In addition to finding that agreement was low to moderate, parents consistently reported lower overall and domain WHODAS 2.0 scores compared to youth. The direction of parent-youth discrepancies is inconsistent in literature. Similar to the results of this thesis, a school-based sample found that parents underestimated children and adolescent's emotional distress symptoms (92). Multiple reviews however, focused on children with health problems, have identified that parents tend to report worse quality of life compared to children themselves (33,93). In a related vein, it may be expected that parents are likely to report higher functional impairment compared to youth. In fact, a study of youth receiving outpatient mental health services found that parents reported increased levels of functional impairment compared to youth reports (94). This sample was predominantly ethnic minority youth (94), while the majority of youth considered in this thesis were Caucasian. Based on the conflicting results, it is possible that the direction of informant discrepancies may be explained, in part, by ethnicity. Previous work has identified ethnicity as a significant predictor of the magnitude and direction of informant discrepancies (95–98). Parents from ethnic minorities may have unique thresholds, compared to children and non-minority

parents, at determining when medical attention for poor mental health is needed (98,99). As culture impacts the conceptualization of family, it is reasonable to expect that parent-child communication and ultimately informant discrepancies may differ across ethnic groups (100). Parents of ethnic minorities also may be less likely to endorse symptoms of mental disorders due to perceived stigma (101). Overall, additional research is needed with more diverse samples such that the impact of ethnicity can be assessed.

As agreement tends to be higher for characteristics that can be observed (39,44), it was expected that the mobility and self-care domain scores would have the highest agreement compared to the other WHODAS 2.0 domain scores. When considering the Bland-Altman plots for self-care and mobility domain scores, variability between parent and youth responses increased as the average WHODAS 2.0 score increased. The increased variation may suggest that as functional impairment increases, the magnitude of the difference between parent and youth reports also increases. Consistent with the hypothesis, the domain with the greatest agreement was the self-care domain. However, agreement on the mobility domain was lower than agreement on the life activities and participation domains. Agreement tends to increase on domains related to the condition of interest, as frequent communication about that domain may occur (33,35). It is possible that mobility concerns are not a focus of clinical or personal discussions about a youth's mental health and physical functioning. However, mobility can be impacted by mental disorder. For example, individuals with anxiety and depression may have idiopathic clinical symptoms such as pain, which could decrease physical activity (102). As youth are still physically able to complete tasks (i.e. standing and walking), individuals may not consider these declines in physical activity to be the same as limitations in mobility, impacting responses on the mobility domain. Concerns may instead be discussed in the context of low motivation to complete tasks and connect with others, not decreased physical ability, potentially explaining the increased agreement on the life activities and participation domains.

Bland-Altman plots and ICC values both suggested that youth and parents provide discrepant information on the WHODAS 2.0, regardless of the version or domain being considered. As parents and youth provide discrepant reports, information from both informants should be collected and considered when trying to assess functional impairment among youth with mental disorder. While it is ideal to collect information from both informants (35,55,103), this may not always be possible in clinical settings where time and resources may be limited. Unlike child

populations, where there may be a concern that an informant's young age limits their ability to articulate information accurately, it is unlikely that this is true for youth. Therefore, if one informant perspective must be prioritized, it may be best to prioritize the youth perspective. Ultimately, it is the youth who will best understand their internal states (35,104). Placing the focus on youth perspectives is also valuable as youth are expected to become partners in their care, and prepare to transition into adult care settings where they will likely be the sole informant (49).

Bland-Altman plots revealed that parent-youth agreement is similar when considering the 35- and the 12-item version of the WHODAS 2.0. While the 12-item WHODAS 2.0 explains the majority (81%) of the variance on the 36-item WHODAS 2.0 (30), it was important to consider both versions of the WHODAS 2.0 to ensure that agreement between informants didn't differ based on the version of the WHODAS 2.0 used. Most literature exploring agreement on the WHODAS 2.0 considers either the 12-item version or the 36-item version of the WHODAS 2.0. By showing that parent-youth agreement is similar for both versions, the results provide evidence that degree of parent-youth agreement from investigations using only one version may be generalized to both versions of the WHODAS 2.0. Second, by finding that agreement is low to moderate for both versions of the WHODAS 2.0, when assessing the functional impairment of youth with mental disorder, the less time-consuming 12-item version of the WHODAS 2.0 may be preferred in both clinical and research settings.

5.2. Objective 2

Convergent and divergent validity was assessed by considering the effect size of correlations. The hypothesis that KIDSCREEN domains would be correlated with the 12-item and 36-item WHODAS 2.0 scores was not supported, as all correlations were low to moderate. The hypothesis that demographic variables (youth age, youth sex, and household income) would not be associated with 12-item and 36-item WHODAS 2.0 scores was supported, as all correlations were low to moderate. Additionally, the hypothesis that the strength of correlations would be similar between parent and child report was largely supported. The single exception was observed when considering the correlation between the 35-item WHODAS 2.0 scores and the KIDSCREEN *physical well-being* domain score; the correlation between the youth 35-item WHODAS 2.0 score and the youth reported *physical well-being* score was moderate, while the parent reported

correlation was low. Finally, as expected, the internal consistency for both informants was adequate.

Increases in functional impairment are generally associated with decreases in health-related quality of life (105). While the results demonstrate this inverse association, convergent validity of the WHODAS 2.0 was not demonstrated for either informant, as all correlations between WHODAS 2.0 scores and KIDSCREEN domains were low to moderate. This was true when considering both 35-item WHODAS 2.0 scores, and 12-item WHODAS 2.0 scores. It was surprising that KIDSCREEN domain scores and the WHODAS 2.0 overall scores were not strongly correlated. Although functional impairment and health-related quality of life are related, they are distinct constructs (106). The WHODAS 2.0 doesn't emphasize social aspects of disability (107), these aspects of disability may be better captured by the KIDSCREEN, potentially explaining why strong correlations were not observed. Evidence also suggests that youth with chronic conditions report health-related quality of life similar to, or better than, healthy children (33,49). As a subset of the sample included in this work had chronic conditions, it is possible that correlations between the WHODAS 2.0 and the KIDSCREEN were negated. Regardless, the WHODAS 2.0 appeared to be performing similarly between informants as youth reported and parent reported correlations between KIDSCREEN domains and WHODAS 2.0 scores were similar in direction and magnitude.

Divergent validity of the WHODAS 2.0 was demonstrated as there was an absence of strong correlations with demographic variables (female youth, youth age, and annual household income <$75,000). Demographic variables were selected to assess divergent validity as it is not expected that functional impairment will differ with youth age, sex, or household income. These results add further support to literature showing divergent validity of the WHODAS 2.0 has been established in diverse samples (30,64), and identify that this extends to both a clinical sample of youth with mental disorder and their parents. Again, this strength and direction of youth reported, and parent reported correlations were comparable.

In addition to the findings that youth and parent-reported correlations (for both convergent and divergent validity) are similar, the internal consistency of both informants was adequate. These reports align with other investigations of the internal consistency of the WHODAS 2.0 in youth, and in those with various mental disorders (26,31). As informant age decreases there is increased fear that an individual may not be able to accurately articulate their concerns. By finding that the

psychometric properties of the WHODAS 2.0 were similar for both youth with mental disorder and their parents, these concerns should be minimized. This result supports literature which suggests that the WHODAS 2.0 can be used in youth as young as 15 years old (29). Overall, it seems that both youth with mental disorder and their parents are able to provide valid and reliable reports of functional impairment on the WHODAS 2.0.

5.3. Objective 3

Contrary to the hypothesis, that demographic, youth health, parent health, and healthcare setting covariates would be useful in explaining parent-youth disagreement on the WHODAS 2.0, few covariates were meaningful. Additionally, the hypothesis that the covariates which best explained the variation between parent-youth disagreement on the 35-item version of the WHODAS 2.0 would be the same covariates that explained disagreement on the 12-item version was not supported. In both models, only 1 covariate was significantly associated with meaningful disagreement between parent and youth WHODAS 2.0 scores. When considering the 35-item WHODAS 2.0 scores, household income <$75,000 was associated with lower odds of meaningful differences between parent and youth WHODAS 2.0 scores. This association was observed in a crude model, and when controlling for demographic, youth health, and healthcare setting covariates. When considering the 12-item WHODAS 2.0 scores, higher youth age was associated with lower odds of meaningful differences between parent and youth WHODAS 2.0 scores. This association was also observed in a crude model, and when controlling for demographic, youth health, parent health, and healthcare setting covariates.

The results that showed household income <$75,000 was associated with lower odds of meaningful disagreement between parent and youth 35-item WHODAS 2.0 scores conflicted with other literature, which showed that agreement increases as household income increases (108). The perceived need for mental health supports decreases as household income declines (5). Further, children in low-income households, especially youth, have higher rates of mental illness and the highest unmet need for mental health support (8,109). A myriad of factors – from low health literacy to decreased time for leisure – may contribute to these associations. Knowing this, potential explanations, for the finding that low household income was associated with decreased odds of meaningful disagreement, emerge. First, the process of overcoming barriers may result in a shared understanding about mental health concerns, increasing the odds of agreement for

informants in low-income households. Second, it is possible that only dyads with the greatest communication between youth and parents may seek mental health services, leading to selection bias within this sample. Thus, the association observed between low household income and disagreement may be better explained by other confounding factors, such as parent-youth relationship and/or communication. These factors were not assessed within this sample but should be considered in future research. It is also plausible that the youth from low-income households included within this sample represent those with the most noticeable functional impairment. If this was true increased agreement would be expected, but this association may disappear if a representative group of youth with mental disorder from low-income households were included. However, these explanations are speculative and further research is required in larger and more socioeconomically diverse samples.

Older youth age was associated with lower odds of meaningful differences between parent and youth 12-item WHODAS 2.0 scores. This result aligns with the literature which suggests that agreement increases as children age (41,46,49). However, the literature is inconclusive, as some studies report the opposite (39,89), and others find no association (48,110). It is difficult to reconcile the discrepant reports about the impact of age, as these studies are completed in diverse samples and inconsistent definitions of 'child' and 'youth' are used. Regardless, it is interesting that an association between youth age and agreement was identified despite the narrow age range of youth within our sample. This finding suggests that even small changes in youth age may be result in meaningful differences in agreement which supports the recommendation that narrow age groupings be used in future research (111). It is possible that, as youth age they are better able and willing to communicate concerns to parent informants. However, additional research should be completed to better understand the association between agreement and youth age.

It was surprising that the regression results differed between WHODAS 2.0 versions, as the 12-item WHODAS 2.0 explains 81% of the variance in the complete 36-item WHODAS 2.0 (30). While the 12-item WHODAS 2.0 has an increased clinical utility, as it takes less time to complete, there is still approximately one fifth of variance in functional impairment not captured on the 12-item version. This additional variance likely clarifies why covariates identified, through regression, differed depending on the WHODAS 2.0 version considered. It is important to recognize that the factors which explain agreement between youth with mental disorder and their parents on one version of the WHODAS 2.0 may not be generalizable to other WHODAS 2.0

versions, as most research to date focuses on only one version of the WHODAS 2.0. Inconsistencies seen between studies exploring factors that may influence agreement on the WHODAS 2.0 may, in part, be explained by the version used.

Associations were not found for youth sex, multimorbidity, youth stress, parent psychosocial distress, or recruitment setting despite literature suggesting that these factors may be important at explaining disagreement between parent and youth informants (32,34,35,42,46,49,51). Our results indicated this may not be true when assessing parent-youth agreement on functional impairment. Current literature is inconsistent about the presence or direction of associations for these covariates. For example, despite both studies assessing parent-youth agreement for anxiety within clinical samples, youth sex was meaningful in one investigation of parent-youth agreement (32), and unimportant in another (46). This thesis took a conservative approach by controlling for factors that have been identified in the literature, even if others negate a potential association. Discrepancies between the results of this thesis and other investigations may also exist; as youth considered in this thesis were older and from a more narrow age range (32,34,35,42,46,49,51). Further, youth included within this sample had various mental disorders; though, the factors that influence agreement are known to differ depending on the specific mental disorder considered (46). It is possible that a factor may increase odds of agreement for one metal disorder but decrease odds of agreement for another mental disorder – thereby nullifying the impact of that variable within this investigation. Additional research should be completed with larger and more diverse samples to elucidate the impact of these covariates on parent-youth agreement.

5.4. Limitations

The results of this thesis need to be considered in the context of several limitations. First, the sample is both small and homogenous. Participants were recruited from a single site, over three-quarters of the sample of youth were female, and almost all were Caucasian. Therefore, the results may not have been powered to detect significant associations based on sex or ethnicity, and are not generalizable to all youth with mental disorder. Further, the sample is underpowered to examine illness-specific agreement. Research reports that agreement is consistent across diagnoses (90). However, there is evidence that parent-child agreement differs based on illness type (i.e. internalizing vs. externalizing), and that agreement differs depending on the specific mental

disorder considered (44,46). Regardless, it must be remembered that the results of this thesis are preliminary; more investigations should be completed with clinical samples of youth that independently consider each mental disorder. Third while the parent informants considered were the primary caregiver, this investigation did not consider the gender of the parent informant. Similar to other literature, which finds that the parent informant tends to be the mother (112), most of parent informants in this work were female. However, parent informants are not interchangeable as mother-youth agreement may not be the same as father-youth agreement (113). Thus, the results of this thesis may not be generalizable to father-youth dyads. Future work should include larger samples and place an emphasis on collecting information from both parents, where possible. Moreover, this dataset did have some missing data. The use of mean imputation was valuable to maximize the number of participants that could be included within this study. However, imputation methods reduced the standard deviation of the dataset. As is common with all research, future studies should attempt to minimize missing responses and maximize participation. Next, a clinically meaningful important difference score for the WHODAS 2.0 had not yet been determined (26). While the methods used to classify meaningful disagreement was consistent with other investigations (87), it is unclear if 0.5 standard deviations was an appropriate threshold. Additional work should be completed to determine the clinically meaningful important difference score for the WHODAS 2.0. Finally, while the simple scoring method was employed in this work, as it is the most likely method to be used in clinical settings, it is unclear if the results would change if the complex scoring method was used. Researchers should be mindful not to generalize these results to studies using WHODAS 2.0 scores generated using the complex scoring method.

5.5. Implications & Future Directions

This thesis has several important implications for research, policy, and practice. By identifying that parent-youth agreement was similar for both the 12-item and 35-item WHODAS 2.0, showing that convergent and divergent validity were similar for both informants, and revealing that internal consistency was sufficient for both youth and parent informants, it is wise to suggest that use of the 12-item WHODAS 2.0 version be prioritized. As the 12-item WHODAS 2.0 is less time-consuming than the longer and more detailed 35-item WHODAS 2.0, this recommendation will be valuable in clinical settings where time may be limited.

It is reasonable to expect that a clinical the sample would have increased agreement, as parents and youth are both aware that clinical support is being received for mental health concerns. However, the results of this study suggest that this is not the case and that low-to-moderate agreement between parents and youth may also extend to clinical samples. By showing that there was low-to-moderate agreement between parent and youth informants, this thesis supports the notion that these informants provide unique and valuable perspectives (40,43,53). Similar to other studies, these findings support that information from both informants should be collected in both clinical and research settings (42,112). Identifying, and hopefully reconciling, both perspectives may better allow health professionals to provide optimal supports and care to youth experiencing mental disorder. Additionally, parent-youth disagreement itself may predict poor outcomes beyond those explained by increased functional impairment (53). Thus, collecting information from both informants may help health professionals identify parent-youth dyads at risk for poor outcomes, and provide supports to reduce discrepancies thereby improving youth health trajectories.

Despite using a narrow age range, results still showed that youth age was associated with meaningful disagreement on the WHODAS 2.0. When it is not possible to collect information from all youth and parent dyads, collecting information from both informants for dyads with young youth should be prioritized as young youth have greater odds of meaningful disagreement on the 12-item WHODAS 2.0. However, given the negative impact that informant discrepancies may have on health outcomes for youth (53), and value of family-centered approaches to care, creating health policies that incentivize collecting information from both informants for all parent-youth dyads may be justified.

While logistic regression results began to elucidate factors that may be associated with meaningful parent-youth disagreement on the WHODAS 2.0, this thesis is not able to explain why disagreement exists. Much of the research currently exploring parent-youth agreement, similar to this investigation, is quantitative. Additional research is needed to assess the reasons why discrepancies arise, not just the factors associated with informant discrepancies (40). Qualitative research methods may be uniquely equipped to answer research questions that could fill this research gap.

Conclusion

Overall, the conclusions derived from both versions of the WHODAS 2.0 are similar. While both versions of the WHODAS 2.0 may have value depending on the purpose of an

assessment, this recognition is valuable for clinical settings where time is often limited, as the 12-item version of the WHODAS 2.0 is less time consuming to complete. Reports from both youth and parents appear valuable in understanding functional impairment. Thus, it is important that the perspectives of both informants are collected and considered in both clinical and research settings. Additional work is needed to understand the factors that influence discrepancies and the implications for care. Preliminary evidence from this thesis suggests that household income and youth age may be valuable in explaining meaningful disagreement between parents and youth. Further investigations should be completed to assess these potential associations in larger and more diverse samples.

REFERENCES

1. Merikangas KR, Nakamura EF, Kessler RC. Epidemiology of mental disorders in children and adolescents. Dialogues Clin Neurosci. 2009;11(1):7–20.
2. Polanczyk G V, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry. 2015;56(3):345–65.
3. Georgiades K, Duncan L, Wang L, Comeau J, Boyle MH. Six-Month Prevalence of Mental Disorders and Service Contacts among Children and Youth in Ontario: Evidence from the 2014 Ontario Child Health Study. Can J Psychiatry. 2019;64(4):246–55.
4. Findlay LC, Sunderland A. Professional and informal mental health support reported by Canadians aged 15 to 24. Ottawa, Canada; 2014.
5. Comeau J, Georgiades K. Changes in the Prevalence of Child and Youth Mental Disorders and Perceived Need for Professional Help between 1983 and 2014: Evidence from the Ontario Child Health Study. Can J Psychiatry. 2019;64(4):257–64.
6. Gandhi S, Chiu M, Lam K, Cairney JC, Guttmann A, Kurdyak P. Mental Health Service Use Among Children and Youth in Ontario: Population-Based Trends Over Time. Can J Psychiatry. 2016;61(2):119–24.
7. Costello JE, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and Development of Psychiatric Disorders in Childhood and Adolescence. Arch Gen Psychiatry. 2003;60:837–44.
8. Lawrence D, Johnson S, Hafekost J, Haan KB de, Sawyer M, Ainley J, et al. The Mental Health of Children and Adolescents. Canberra; 2015.
9. Harwood HJ, Mark TL, Mckusick DR, Coffey RM, King EC, Genuardi JS. National Spending on Mental Health and Substance Abuse Treatment by Age of Clients, 1997. J Behav Health Serv Res. 1997;30(4):433–43.
10. Making the Case for Investing in Mental Health in Canada. Ottawa, Canada; 2013.
11. CMHO. Annual Report Card: The burden of kids mental illness on families and the economy. 2019.
12. Costello JE, Egger H, Angold A. 10-Year Research Update Review: The Epidemiology of Child and Adolescent Psychiatric Disorders: I. Methods and Public Health Burden. J Am Acad Child Adolesc Psychiatry [Internet]. 2005;44(10):972–86. Available from:

http://dx.doi.org/10.1097/01.chi.0000172552.41596.6f
13. Patel V, Flisher AJ, Hetrick S, Mcgorry P. Mental health of young people: a global public-health challenge. Lancet. 2007;369:1302–13.
14. Chen H, Cohen P, Kasen S, Johnson JG, Berenson K, Gordon K. Impact of Adolescent Mental Disorders and Physical Illnesses on Quality of Life 17 Years Later. Arch Pediatr Adolesc Med. 2006;160:93–9.
15. Evans S, Banerjee S, Leese M, Huxley P. The impact of mental illness on quality of life: A comparison of severe mental illness, common mental disorder and healthy population samples. Qual Life Res. 2007;16:17–29.
16. Newacheck PW, Taylor WR. Childhood Chronic Illness: Prevalence, Severity, and Impact. Am J Public Health. 1992;82(3).
17. Hansen MS, Fink P, Frydenberg M, Oxhoj M-L, Sondergaard L, Eriksen M. Mental Disorders in Medical Inpatients and the Association to Severity of Illness, Self-Rated Physical Disability, and Health Perception. Psychosomatics. 2001;41(1):41–7.
18. Suryavanshi MS, Yang Y. Clinical and Economic Burden of Mental Disorders Among Children With Chronic Physical Conditions, United States, 2008–2013. Prev Chronic Dis. 2016;13(150535):1–10.
19. Varni JW, Burwinkle TM, Lane MM. Health-related quality of life measurement in pediatric clinical practice: An appraisal and precept for future research and application. Health Qual Life Outcomes. 2005;3(34):1–9.
20. McDowell I. Measuring health: a guide to rating scales and questionnaires. Third. New York, New York: Oxford University Press; 2006.
21. Iezzoni LI, Ash AS, Coffman GA, Moskowitz MA. Predicting In-Hospital Mortality: A Comparison of Severity Measurement Approaches. Med Care. 1992;30(4):347–59.
22. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. Fifth. Oxford: Oxford University Press; 2015.
23. Hafetz J, Miller VA. Child and Parent Perceptions of Monitoring in Chronic Illness Management: A Qualitative Study. Child Care Heal Dev. 2011;36(5):655–62.
24. Von Korff M, Glasgow RE, Sharpe M. Organising care for chronic illness. BMJ. 2002;325:92–4.
25. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome

monitoring in clinical practice: Benefits, challenges, and solutions. Psychother Res [Internet]. 2015;25(1):6–19. Available from: http://dx.doi.org/10.1080/10503307.2013.817696

26. Federici S, Bracalenti M, Meloni F, Luciano J V. World Health Organization disability assessment schedule 2.0: An international systematic review. Disabil Rehabil. 2017;39(23):2347–80.
27. Goldstein S, Naglieri JA. Assessing Impairment: From Theory to Practice. Second. New York, New York: Springer; 2016.
28. Tompke BK, Tang J, Oltean II, Buchan MC, Reaume S V, Ferro MA. Measurement Invariance of the WHODAS 2.0 Across Youth With and Without Physical or Mental Conditions. Assessment. 2018;1–12.
29. Kimber M, Rehm J, Ferro MA. Measurement Invariance of the WHODAS 2.0 in a Population-Based Sample of Youth. PLoS One [Internet]. 2015;10(11):1–13. Available from: http://dx.doi.org/10.1371/journal.pone.0142385
30. Üstün TB, Kostanjsek N, Chatterji S, Rehm J. Measuring Health and Disability Manual for WHO Disability Assessment Schedule. Geneva, Switzerland; 2010.
31. Hu L, Zang Y, Li N. The applicability of WHODAS 2.0 in adolescents in China. J Clin Nurs. 2012;21:2438–51.
32. Choudhury MS, Pimentel SS, Kendall PC. Childhood Anxiety Disorders: Parent–Child (Dis)Agreement Using a Structured Interview for the DSM-IV. J Am Acad Child Adolesc Psychiatry [Internet]. 2003;42(8):957–64. Available from: http://dx.doi.org/10.1097/01.CHI.0000046898.27264.A2
33. Upton P, Lawford J, Eiser C. Parent–child agreement across child health-related quality of life instruments: a review of the literature. Qual Life Res. 2008;17:895–913.
34. Bagheri Z, Jafari P, Tashakor E, Kouhpayeh A, Riazi H. Assessing Whether Measurement Invariance of the KIDSCREEN-27 across Child-Parent Dyad Depends on the Child Gender: A Multiple Group Confirmatory Factor Analysis. Glob J Health Sci. 2014;6(5):142–53.
35. Eiser C, Varni JW. Health-related quality of life and symptom reporting: similarities and differences between children and their parents. Eur J Pediatr. 2013;172:1299–304.
36. Gregorich SE. Do Self-Report Instruments Allow Meaningful Comparisons Across

Diverse Population Groups? Med Care. 2006;44(11):s78–94.

37. Chang P, Yeh C. Agreement between child self-report and parent proxy-report to evaluate quality of life in children with cancer. Psychooncology. 2005;14:125–34.
38. Dudley N, Ackerman A, Brown KM, Snow SK. Patient- and Family-Centered Care of Children in the Emergency Department. Pediatrics. 2015;135(1).
39. Achenbach TM, Mcconaughy SH, Howell CT. Child/Adolescent Behavioral and Emotional Problems: Implications of Cross-Informant Correlations for Situational Specificity. Psychol Bull. 1987;101(2):213–32.
40. De Los Reyes A, Kazdin AE. Informant Discrepancies in the Assessment of Childhood Psychopathology: A Critical Review, Theoretical Framework, and Recommendations for Further Study. Psychol Bull. 2005;131(4):483–509.
41. Edelbrock C, Costello A, Dulcan M, Conover N, Kala R. Parent-child agreement on child psychiatric symptoms assessed via structured interview. J Child Psychol Psychiatry. 1986;27(2):181–90.
42. Martin JL, Ford CB, Dyer-Friedman J, Tang J, Huffman LC. Patterns of Agreement Between Parent and Child Ratings of Emotional and Behavioral Problems in an Outpatient Clinical Setting: When Children Endorse More Problems. Dev Behav Pediatr. 2004;25(3):11–3.
43. Yeh M, Weisz JR. Why Are We Here at the Clinic? Parent-Child (Dis) Agreement on Referral Problems at Outpatient Treatment Entry. J Consult Clin Psychol. 2001;69(6):1018–25.
44. De Los Reyes A, Augenstein TM, Wang M, Thomas SA, Drabick DAG, Burgers DE, et al. The Validity of the Multi-Informant Approach to Assessing Child and Adolescent Mental Health. Psychol Bull. 2015;141(4):858–900.
45. Hamblin RJ, Salloum A, Andel R, Nadeau JM, Mcbride NM, Lewin AB, et al. Predictors of parent-child agreement on child anxiety diagnoses on the ADIS-IV-C/P. Psychiatry Res [Internet]. 2016;245:303–10. Available from: http://dx.doi.org/10.1016/j.psychres.2016.07.041
46. Becker EM, Jensen-Doss A, Kendall PC, Birmaher B, Ginsburg GS. All Anxiety is not Created Equal: Correlates of Parent/Youth Agreement Vary Across Subtypes of Anxiety. J Psychopathol Behav Assess [Internet]. 2016;38:528–37. Available from:

http://dx.doi.org/10.1007/s10862-016-9544-z

47. Barbosa J, Tannock R, Manassis K. Measuring anxiety: Parent-child reporting differences in clinical samples. Depress Anxiety. 2002;15:61–5.
48. Popp L, Neuschwander M, Mannstadt S, In-albon T, Schneider S. Parent-Child Diagnostic Agreement on Anxiety Symptoms with a Structured Diagnostic Interview for Mental Disorders in Children. Front Psychol. 2017;8:1–12.
49. Sattoe JNT, van Staa A, Moll HA. The proxy problem anatomized: Child-parent disagreement in Health Related Quality of Life reports of chronically ill adolescents. Health Qual Life Outcomes [Internet]. 2012;10(1). Available from: http://www.hqlo.com/content/10/1/10
50. Lee L, Halpern CT, Hertz-Picciotto I, Martin SL, Suchindran CM. Child care and social support modify the association between maternal depressive symptoms and early childhood behaviour problems: a US national study. J Epidemiol Community Heal. 2006;60:305–10.
51. Oltean II, Ferro MA. Agreement of child and parent-proxy reported health-related quality of life in children with mental disorder. Qual Life Res [Internet]. 2019;28(3):703–12. Available from: http://dx.doi.org/10.1007/s11136-018-2026-x
52. Shaw K., Southwood T., McDonagh J. Growing Up and Moving On in Rheumatology: Parents as Proxies of Adolescents With Juvenile Idiopathic Arthritis. Arthritis Rheum. 2006;55(2):189–98.
53. Ferdinand RF, van der Ende J, Verhulst FC. Parent–Adolescent Disagreement Regarding Psychopathology in Adolescents From the General Population as a Risk Factor for Adverse Outcome. J Abnorm Psychol Copyr 2004 by Am Psychol Assoc. 2004;113(2):198–206.
54. Ravelli A, Viola S, Migliavacca D, Pistorio A, Ruperto N, Martini A. Discordance between proxy-reported and observed assessment of functional ability of children with juvenile idiopathic arthritis. Rheumatology. 2001;40:914–9.
55. Lee Y, Yang H, Lee W, Teng M. Do parents and children agree on rating a child's HRQOL? A systematic review and Meta-analysis of comparisons between children with attention deficit hyperactivity disorder and children with typical development using the PedsQL. Disabil Rehabil [Internet]. 2017;41(3):265–75. Available from:

https://doi.org/10.1080/09638288.2017.1391338

56. Ferro MA, Lipman EL, Van Lieshout RJ, Boyle MH, Gorter JW, Macmillan HL, et al. Mental–Physical Multimorbidity in Youth: Associations with Individual, Family, and Health Service Use Outcomes. Child Psychiatry Hum Dev [Internet]. 2019;50:400–10. Available from: http://dx.doi.org/10.1007/s10578-018-0848-y
57. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.
58. Anthoine E, Moret L, Regnault A, Sbille V, Hardouin J-B. Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures. Health Qual Life Outcomes. 2014;176(12):1–10.
59. Shrout PE. Measurement reliability and agreement in psychiatry. Stat Methods Med Res. 1998;7:301–17.
60. Hertzog MA. Considerations in Determining Sample Size for Pilot Studies. Res Nurs Health. 2008;31:180–91.
61. Sheehan D V, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, et al. Reliability and Validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). US Dep Heal Hum Serv. 2010;71(3).
62. Duncan L, Georgiades K, Wang L, Lieshout RJ Van, Macmillan HL, Ferro MA, et al. Psychometric Evaluation of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). Psychol Assess. 2018;30(7):916–28.
63. Rehm J, Üstün TB, Saxena S, Nelson CB, Chatterji S. On the development and psychometric testing of the WHO screening instrument to assess disablement in the general population. Int J Methods Psychiatr Res. 1999;8(2):110–22.
64. Saltychev M, Katajapuu N, Bärlund E, Laimi K. Psychometric properties of 12-item self-administered World Health Organization disability general population and people with non-acute physical causes of disability – systematic review Psychometric properties of 12-item self-administered World Health Orga. Disabil Rehabil [Internet]. 2019;1–6. Available from: https://doi.org/10.1080/09638288.2019.1643416
65. Robitail S, Ravens-Sieberer U, Simeoni MC, Rajmil L, Bruil J, Power M, et al. Testing the structural and cross-cultural validity of the KIDSCREEN-27 quality of life

questionnaire. Qual Life Res. 2007;16:1335–45.
66. Tompke BK, Ferro MA. Measurement Invariance and Informant Discrepancies of the KIDSCREEN-27 in Children with Mental Disorder. Appl Res Qual Life. 2019;11:1–20.
67. Ravens-Sieberer U, Auquier P, Erhart M, Gosch A, Rajmil L, Bruil J, et al. The KIDSCREEN-27 quality of life measure for children and adolescents: psychometric results from a cross-cultural survey in 13 European countries. Qual Life Res. 2007;16:1347–56.
68. Cohen S. Perceived Stress Scale. Mind Garden; 1994.
69. Cohen S, Kamarck T, Mermelstein R. A Global Measure of Perceived Stress. J Health Soc Behav. 1983;24(4):385–96.
70. Chiu Y, Lu FJ, Lin J, Nien C, Hsu Y, Liu H. Psychometric properties of the Perceived Stress Scale (PSS): measurement invariance between athletes and non- athletes and construct validity. PeerJ. 2016;
71. Dol M, Mcdonald E, Ferro MA. Psychometric properties of the CESD, STAI-T, and PSS among parents of children with mental illness. J Fam Stud [Internet]. 2020;1–17. Available from: https://doi.org/10.1080/13229400.2020.1837208
72. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl Psychol Meas. 1977;1(3):385–401.
73. Roth DL, Ackerman ML, Okonkwo OC, Burgio LD. The Four-Factor Model of Depressive Symptoms in Dementia Caregivers: A Structural Equation Model of Ethnic Differences. Psychol Aging. 2008;23(3):567–76.
74. Knight RG, Williams S, McGee R, Olaman S. Psychometric properties of the Centre for Epidemiologic Studies Depression Scale (CES-D) in a sample of women in middle life. Behav Res Ther. 1997;35(4):373–80.
75. Devins GM, Orme CM, Costello CG, Binik YM, Frizzell B, Stam HJ, et al. Measuring depressive symptoms in illness populations: Psychometric properties of the Center for Epidemiologic Studies Depression (CES-D) Scale. Psychol Heal. 1998;2:139–56.
76. Spielberger CD. State-Trait Anxiety Inventory for Adults. Tampa, Florida; 1983.
77. Julian LJ. Measures of Anxiety. Arthritis Care Res (Hoboken). 2011;63(S11):S467–72.
78. Barnes LL., Harp D, Jung WS. Reliability generalization of scores on the Spielberger state-trait anxiety inventory. Educ Psychol Meas. 2002;62(4):603–18.

79. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;327(8):307–10.
80. Lu M-J, Zhong Y-X, Miao H-Z, Li Y-C, Ji M-H. Sample Size for Assessing Agreement between Two Methods of Measurement by Bland−Altman Method. Int J Biostat. 2016;
81. Shrout PE, Fleiss JL. Intraclass Correlations: Uses in Assessing Rater Reliability. Psychol Bull. 1979;86(2):420–8.
82. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med [Internet]. 2016;15:155–63. Available from: http://dx.doi.org/10.1016/j.jcm.2016.02.012
83. Ghent AW. Examination of Five Tau Variants Suited to Ordered Contingency Tables, from the Viewpoint of Biological Research. Am Midl Nat. 1984;112(2):332–48.
84. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Second. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
85. Gadermann AM, Zumbo BD, Gadermann AM. Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. Pract Assessment, Res Eval. 2012;17(5).
86. Streiner DL. A Checklist for Evaluating the Usefulness of Rating Scales*. Res Methods Psychiatry. 1993;38:140–8.
87. Qadeer RA, Ferro MA. Child–parent agreement on health-related quality of life in children with newly diagnosed chronic health conditions: a longitudinal study. Int J Adolesc Youth [Internet]. 2018;23(1):99–108. Available from: http://dx.doi.org/10.1080/02673843.2017.1297242
88. Caetano S., Sonpavde G. C-statistic : A brief explanation of its construction, interpretation and limitations. Eur J Cancer. 2018;90:130–2.
89. Rajmil L, López AR, López-Aguilà S, Alonso J. Parent–child agreement on health-related quality of life (HRQOL): a longitudinal study. Health Qual Life Outcomes. 2013;11(101):1–10.
90. Jensen PS, Rubio-Stipec M, Canino G, Bird HR, Dulcan MK, Schwab-Stone ME, et al. Parent and Child Contributions to Diagnosis of Mental Disorder: Are Both Informants Always Necessary? J Am Acad Child Adolesc Psychiatry [Internet]. 1999;38(12):1569–79. Available from: http://dx.doi.org/10.1097/00004583-199912000-00019

91. Weems CF, Feaster DJ, Horigian VE, Robbins MS. Parent and Child Agreement on Anxiety Disorder Symptoms Using the DISC Predictive Scales. Assessment. 2011;18(2):213–6.
92. Kim C, Choi H, Ko H, Park CG. Agreement Between Parent Proxy Reports and Self-Reports of Adolescent Emotional Distress. J Sch Nurs. 2020;36(2):104–11.
93. Galloway H, Newman E. Is there a difference between child self-ratings and parent proxy-ratings of the quality of life of children with a diagnosis of attention-deficit hyperactivity disorder (ADHD)? A systematic review of the literature. ADHD Atten Deficit Hyperact Disord. 2017;9(1):11–29.
94. Yeh M, Zerr A, La R, McCabe K. An Exploration of Parent-Youth Agreement on Functional Impairment in Adolescents Utilizing Outpatient Mental Health Services. Healthcare. 2018;6:1–12.
95. Dirks MA, Weersing VR, Warnick E, Gonzalez A, Alton M, Dauser C, et al. Parent and youth report of youth anxiety: evidence for measurement invariance. J Child Psychol Psychiatry. 2013;55(3):284–91.
96. Youngstrom E, Loeber R, Stouthamer-Loeber M. Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. J Colsulting Clin Psychol. 2000;68(6):1038–1050.
97. Carlston DL, Ogles BM. Age, Gender, and Ethnicity Effects on Parent – Child Discrepancy Using Identical Item Measures. J Child Fam Stud. 2009;18:125–35.
98. Lau AS, Garland AF, Yeh M, McCabe KM, Wood PA, Hough RL. Race/Ethnicity and Inter-Informant Agreement in Assessing Adolescent Psychopathology. J Emot Behav Disord. 2004;12(3):145–56.
99. Roberts RE, Alegria M, Roberts CR, Chen IG. Concordance of Reports of Mental Health Functioning by Adolescents and Their Caregivers: A Comparison of European, African and Latino Americans. J Nerv Ment Dis. 2005;193(8):528–34.
100. Rapp AM, Lau A, Chavira DA. Journal of Anxiety Disorders Differential associations between Social Anxiety Disorder, family cohesion, and suicidality across racial/ethnic groups: Findings from the National Comorbidity Survey-Adolescent (NCS-A). J Anxiety Disord [Internet]. 2016;48:13–21. Available from: http://dx.doi.org/10.1016/j.janxdis.2016.09.009

101. Vega WA, Rodriguez MA, Ang A. Addressing stigma of depression in Latino primary care patients. Gen Hosp Psychiatry. 2010;32(2):182–91.
102. Escobar JI, Dìaz-Martínez A, Gara M. Idiopathic physical symptoms: a common manifestation of psychiatric disorders in primary care. CNS Spectr. 2014;11(3):201–10.
103. Berman AH, Liu B, Ullman S, Jadbäck I, Engström K. Children's Quality of Life Based on the Ratings and Child-Parent Agreement in a Swedish Random Population Sample. PLoS One. 2016;11(3):1–15.
104. Lauth B, Arnkelsson GB, Magnússon P, Skarphéðinsson GÁ, Ferrari P, Pétursson H. Parent–youth agreement on symptoms and diagnosis: Assessment with a diagnostic interview in an adolescent inpatient clinical population. J Physiol - Paris [Internet]. 2010;104(6):315–22. Available from: http://dx.doi.org/10.1016/j.jphysparis.2010.09.004
105. Varni JW, Limbers CA, Burwinkle TM. Parent proxy-report of their children's health-related quality of life: an analysis of 13,878 parents' reliability and validity across age subgroups using the PedsQL ™ 4.0 Generic Core Scales. Health Qual Life Outcomes. 2007;5(2):1–10.
106. Palermo TM, Long AC, Lewandowski AS, Drotar D, Quittner AL, Walker LS. Evidence-based Assessment of Health-related Quality of Life and Functional Impairment in Pediatric Psychology. J Pediatr Psychol. 2008;33(9):983–96.
107. Kulnik ST, Nikoletou D. WHODAS 2.0 in community rehabilitation: a qualitative investigation into the validity of a generic patient-reported measure of disability. Disabil Rehabil. 2014;36(2):146–54.
108. Roy B Van, Groholt B, Heyerdahl S, Clench-aas J. Understanding discrepancies in parent-child reporting of emotional and behavioural problems: Effects of relational and socio-demographic factors. BMC Psychiatry. 2010;56.
109. An R. Unmet mental health care needs in U.S. children with medical complexity, 2005 – 2010. J Psychosom Res [Internet]. 2016;82:1–3. Available from: http://dx.doi.org/10.1016/j.jpsychores.2015.12.007
110. Affrunti NW, Woodruff-Borden J. The effect of maternal psychopathology on parent–child agreement of child anxiety symptoms: A hierarchical linear modeling approach. J Anxiety Disord [Internet]. 2015;32(1):56–65. Available from: http://dx.doi.org/10.1016/j.janxdis.2015.03.010

111. FDA. Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims Guidance for Industry. 2009.
112. Klein RG. Parent-Child Agreement in Clinical Assessment of Anxiety and Other Psychopathology: A Review. J Anxiety Disord. 1991;5:187–98.
113. Eiser C, Havermans T, Pancer M, Eiser JR. Adjustment to Chronic Disease in Relation to Age and Gender: Mothers' and Fathers' Reports of Their Childrens' Behavior. J Pediatr Psychol. 1992;17(3):261–75.

Appendix A – Hypothesized Main-Effects Models

Model 1:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} ;$$
$$\eta_i = \beta_0 + \beta_1 Age_i + \beta_2 Sex_i + \beta_3 Income_i$$

Where:

η_i is the unknown and unobserved odds of a clinically meaningful difference in WHODAS 2.0 score of the *i*-th dyad,
X_{1i} is the observed continuous predictor value for youth age (in years) of the *i*-th subject,
X_{2i} is the observed binary predictor for youth sex (0= female, 1= male) of the *i*-th subject,
X_{3i} is the observed binary predictor for family income (0 ≤ $75000/ year, 1 > $75000/ year) of the *i*-th subject,
β_0 is the fixed unknown intercept,
β_1 is the fixed unknown regression coefficient corresponding to youth age,
β_2 is the fixed unknown regression coefficient corresponding to youth sex,
β_3 is the fixed unknown regression coefficient corresponding to income.

Assumption for any $i \neq j,\ (X_i, Y_i) \perp (X_j, Y_j)$.

Model 2:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} ;$$
$$\eta_i = \beta_0 + \beta_1 Age_i + \beta_2 Sex_i + \beta_3 Income_i + \beta_4 Multimorbidity_i + \beta_5 YouthStress_i$$

Where:

η_i is the unknown and unobserved odds of a clinically meaningful difference in WHODAS 2.0 score of the *i*-th dyad,
X_{1i} is the observed continuous predictor value for youth age (in years) of the *i*-th subject,
X_{2i} is the observed binary predictor for youth sex (0= female, 1= male) of the *i*-th subject,
X_{3i} is the observed binary predictor for family income (0 > $75000/ year, 1 ≤ $75000/ year) of the *i*-th subject,
X_{4i} is the observed binary predictor for youth multimorbidity (No = 0, Yes = 1) of the *i*-th subject,
X_{5i} is the observed continuous predictor value for youth stress of the *i*-th subject,
β_0 is the fixed unknown intercept,
β_1 is the fixed unknown regression coefficient corresponding to youth age,
β_2 is the fixed unknown regression coefficient corresponding to youth sex,
β_3 is the fixed unknown regression coefficient corresponding to income,
β_4 is the fixed unknown regression coefficient corresponding to youth multimorbidity, and
β_5 is the fixed unknown regression coefficient corresponding to youth stress.

Assumption for any $i \neq j,\ (X_i, Y_i) \perp (X_j, Y_j)$.

Model 3:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i};$$
$$\eta_i = \beta_0 + \beta_1 Age_i + \beta_2 Sex_i + \beta_3 Income_i + \beta_4 Multimorbidity_i + \beta_5 YouthStress_i + \beta_6 PsychosocialDistress_i$$

Where:

η_i is the unknown and unobserved odds of a clinically meaningful difference in WHODAS 2.0 score of the *i*-th dyad,
X_{1i} is the observed continuous predictor value for youth age (in years) of the *i*-th subject,
X_{2i} is the observed binary predictor for youth sex (0= female, 1= male) of the *i*-th subject,
X_{3i} is the observed binary predictor for family income (0 > $75000/ year, 1 ≤ $75000/ year) of the *i*-th subject,
X_{4i} is the observed binary predictor for youth multimorbidity (No = 0, Yes = 1) of the *i*-th subject,
X_{5i} is the observed continuous predictor value for youth stress of the *i*-th subject,
X_{6i} is the observed continuous predictor value for parent psychosocial distress of the *i*-th subject,
β_0 is the fixed unknown intercept,
β_1 is the fixed unknown regression coefficient corresponding to youth age,
β_2 is the fixed unknown regression coefficient corresponding to youth sex,
β_3 is the fixed unknown regression coefficient corresponding to income,
β_4 is the fixed unknown regression coefficient corresponding to youth multimorbidity,
β_5 is the fixed unknown regression coefficient corresponding to youth stress, and
β_6 is the fixed unknown regression coefficient corresponding to parent psychosocial distress.

Assumption for any $i \neq j$, $(X_i, Y_i) \perp (X_j, Y_j)$.

Model 4:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i};$$
$$\eta_i = \beta_0 + \beta_1 Age_i + \beta_2 Sex_i + \beta_3 Income_i + \beta_4 Multimorbidity_i + \beta_5 YouthStress_i + \beta_6 PsychosocialDistress_i + \beta_7 Setting_i$$

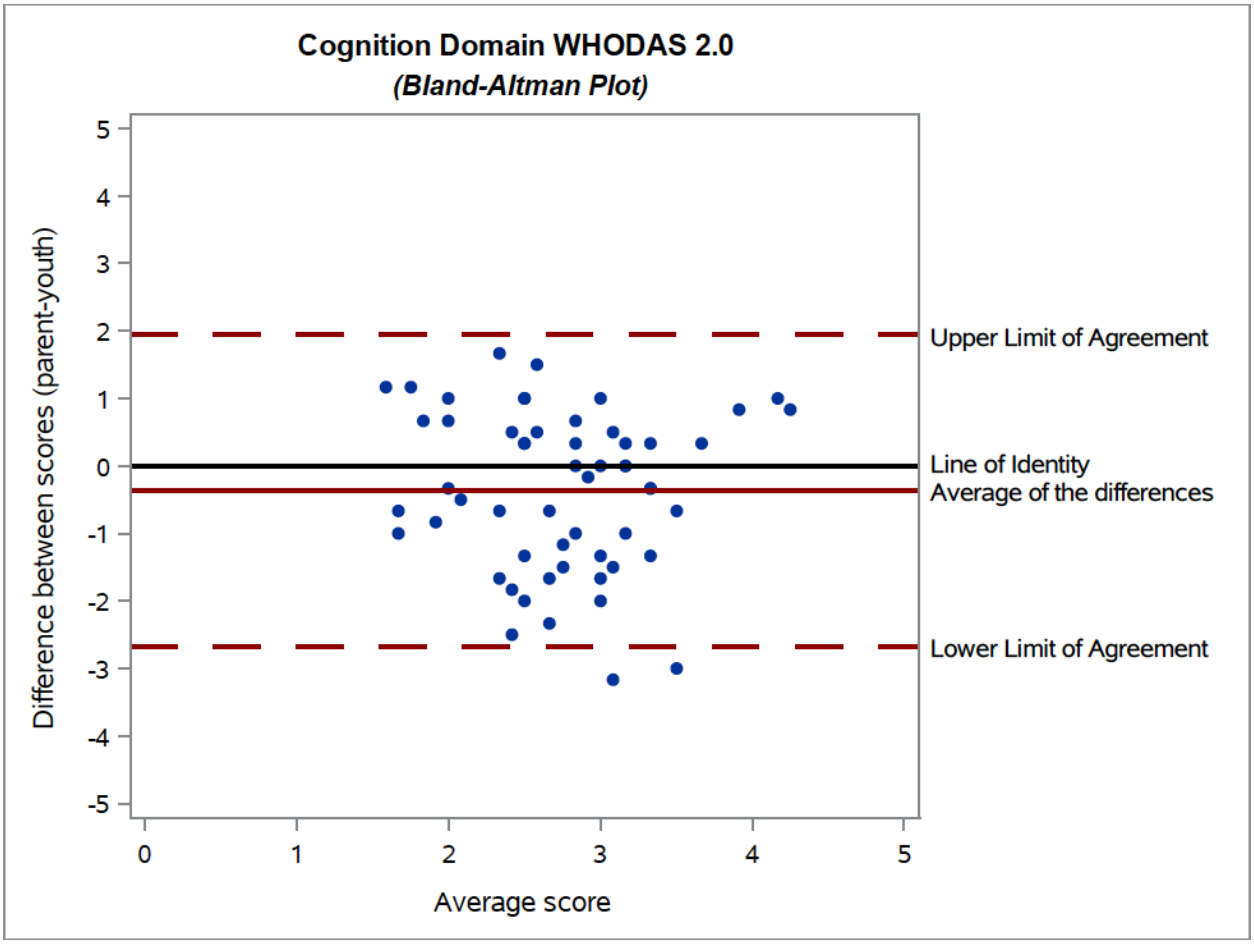
Where:

η_i is the unknown and unobserved odds of a clinically meaningful difference in WHODAS 2.0 score of the *i*-th dyad,
X_{1i} is the observed continuous predictor value for youth age (in years) of the *i*-th subject,
X_{2i} is the observed binary predictor for youth sex (0= female, 1= male) of the *i*-th subject,
X_{3i} is the observed binary predictor for family income (0 > $75000/ year, 1 ≤ $75000/ year) of the *i*-th subject,
X_{4i} is the observed binary predictor for youth multimorbidity (No = 0, Yes = 1) of the *i*-th subject,
X_{5i} is the observed continuous predictor value for youth stress of the *i*-th subject,
X_{6i} is the observed continuous predictor value for parent psychosocial distress of the *i*-th subject,

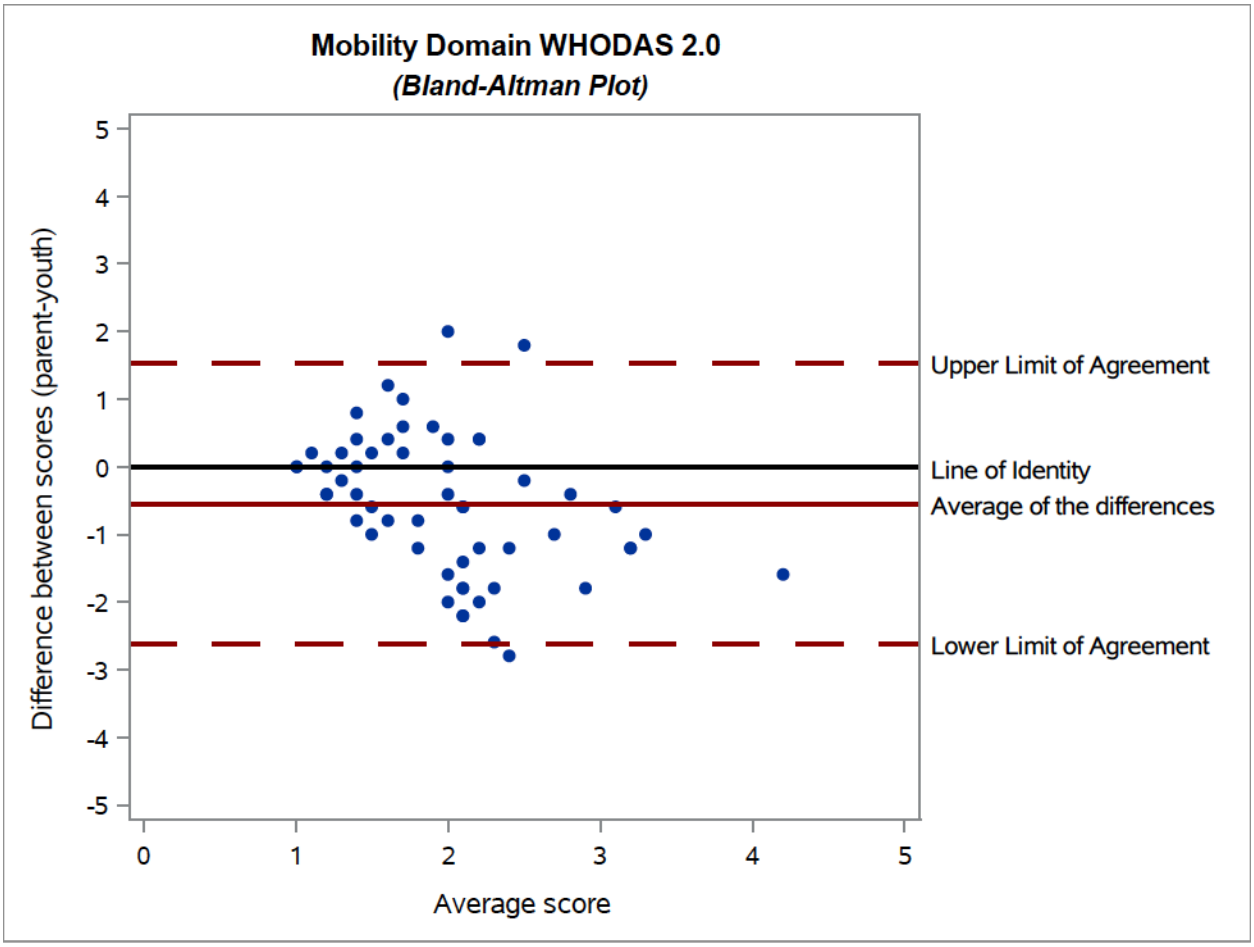
X_{7i} is the observed binary predictor for care setting (0= Inpatient, 1= Outpatient) of the *i*-th subject,
β_0 is the fixed unknown intercept,
β_1 is the fixed unknown regression coefficient corresponding to youth age,
β_2 is the fixed unknown regression coefficient corresponding to youth sex,
β_3 is the fixed unknown regression coefficient corresponding to income,
β_4 is the fixed unknown regression coefficient corresponding to youth multimorbidity,
β_5 is the fixed unknown regression coefficient corresponding to youth stress,
β_6 is the fixed unknown regression coefficient corresponding to parent psychosocial distress, and
β_7 is the fixed unknown regression coefficient corresponding to care setting.

Assumption for any $i \neq j,\ (X_i, Y_i) \perp (X_j, Y_j)$.

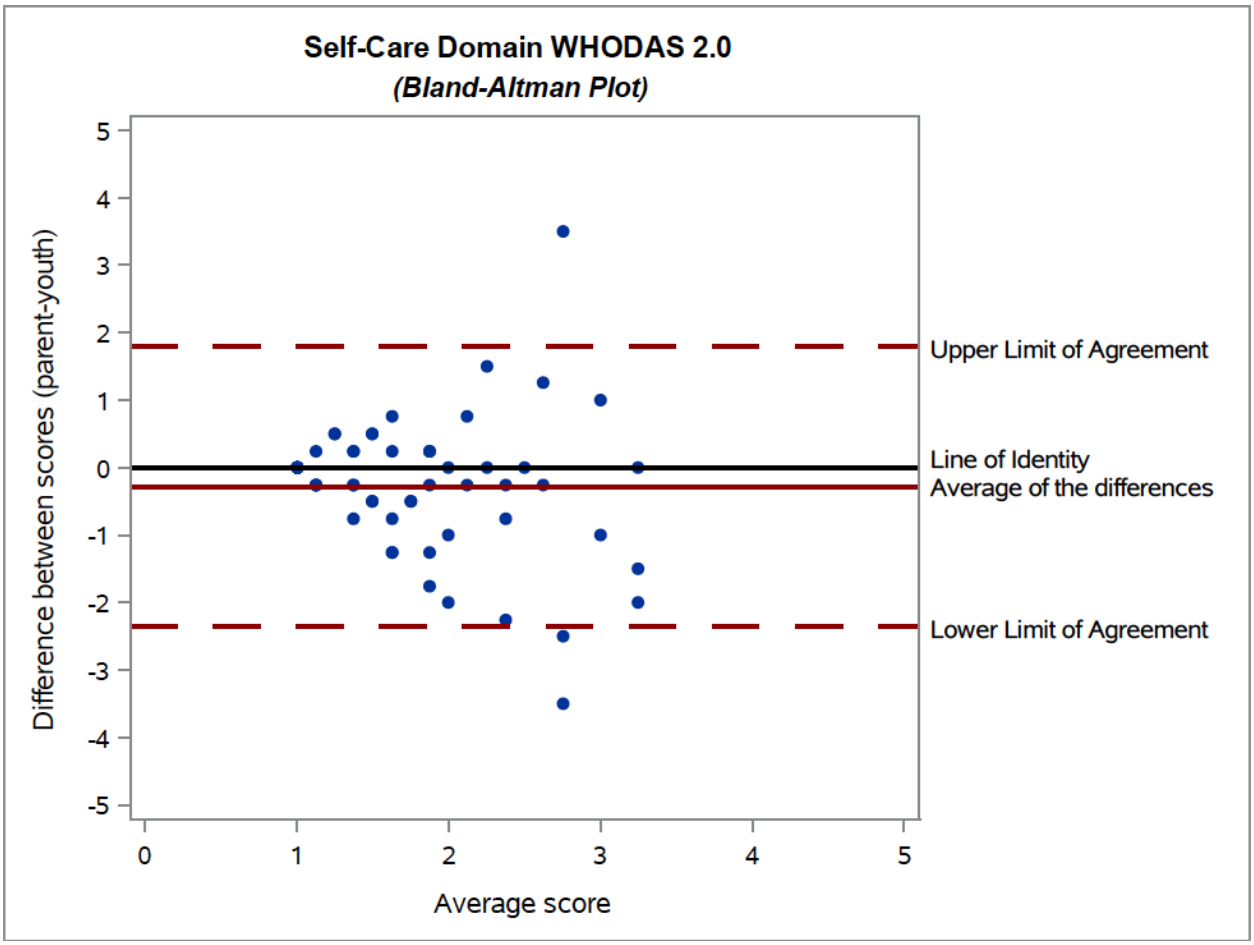
Appendix B – Bland-Altman Plots, WHODAS 2.0 Domain Scores



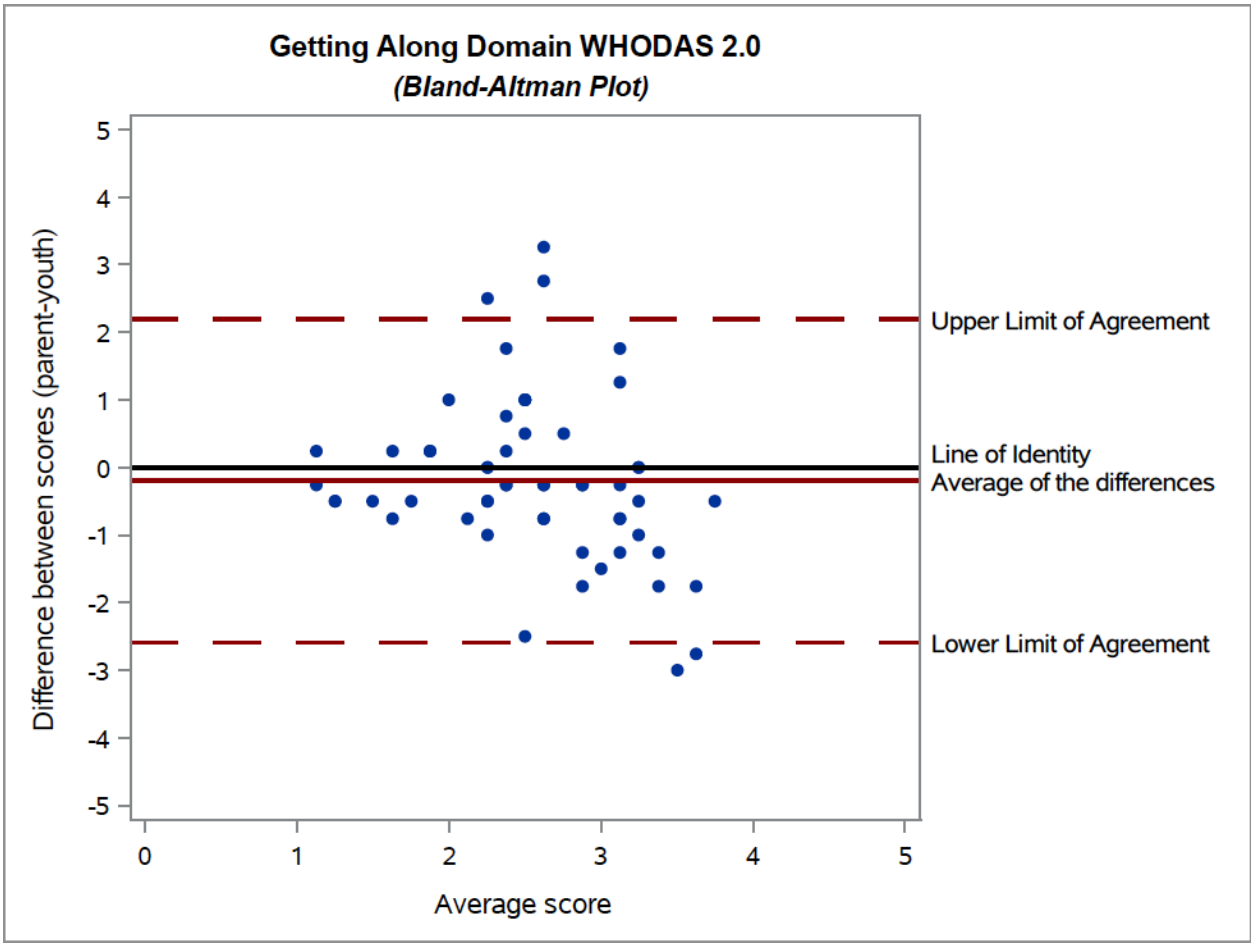
Appendix B – Figure 1. Cognition Domain Bland-Altman Plot



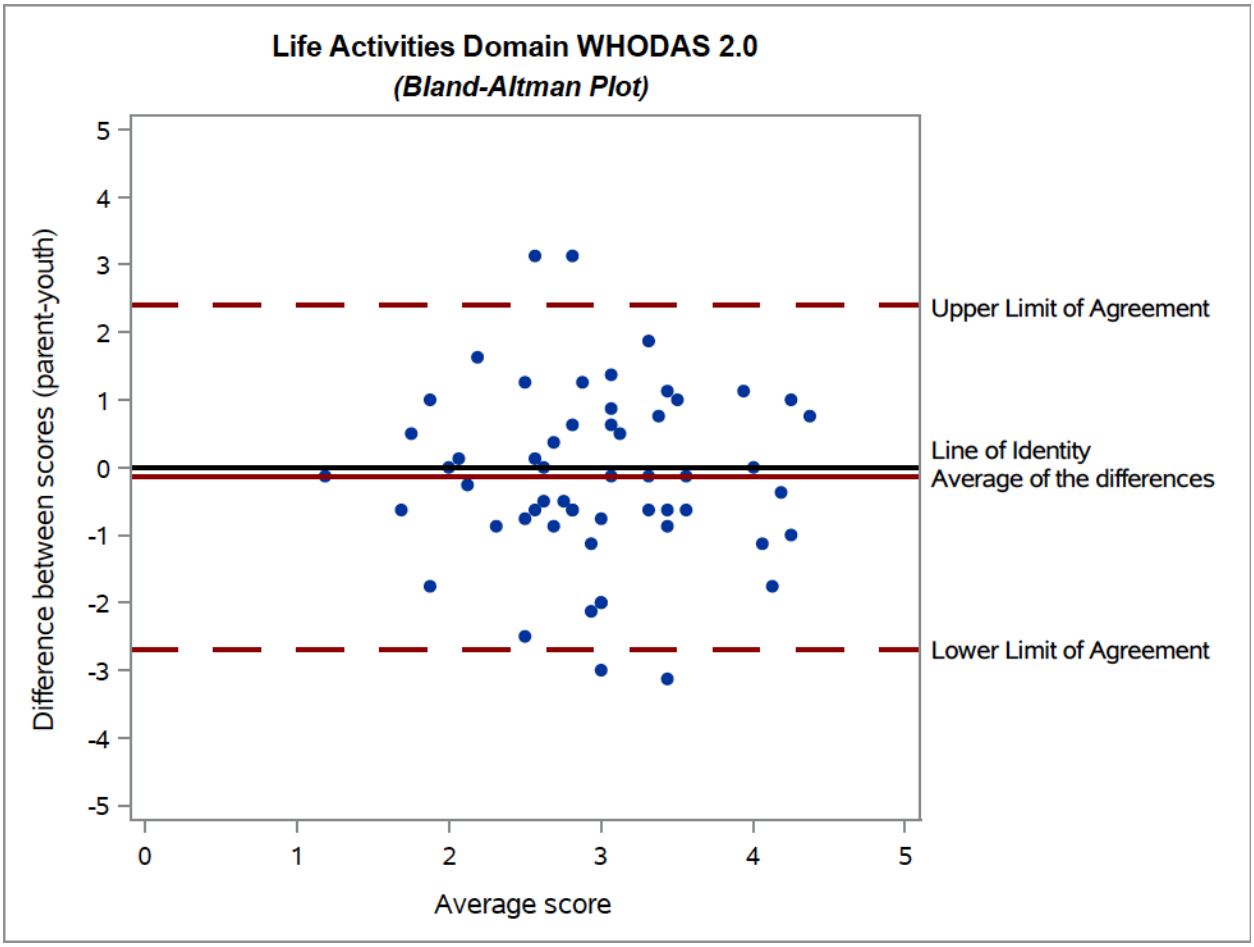
Appendix B – Figure 2. Mobility Domain Bland-Altman Plot



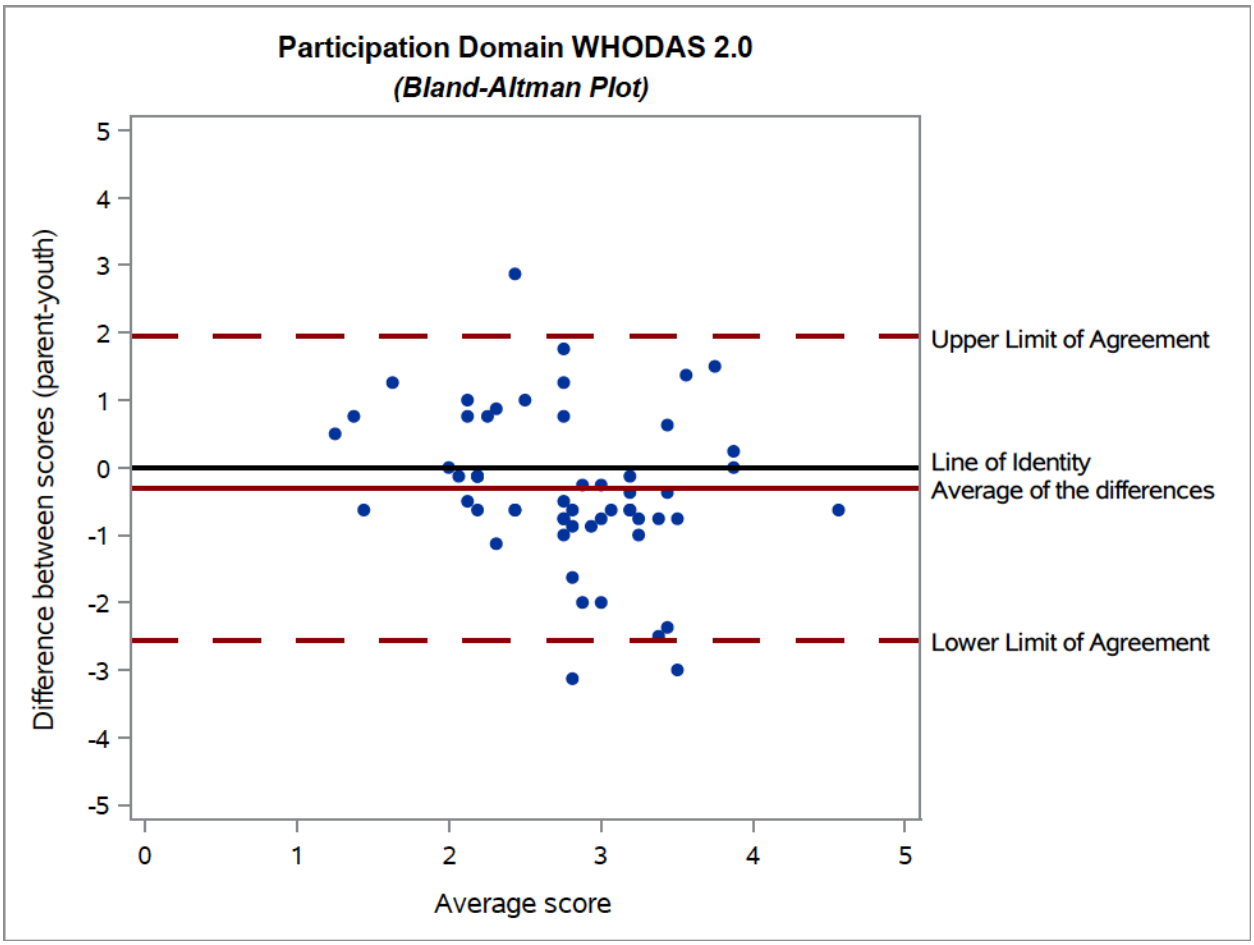
Appendix B – Figure 3. Self-Care Domain Bland-Altman Plot



Appendix B – Figure 4. Getting Along Domain Bland-Altman Plot



Appendix B – Figure 5. Life Activities Domain Bland-Altman Plot



Appendix B – Figure 6. Participation Domain Bland-Altman Plot