Validity of the Flemish Working Alliance Inventory in a Dutch physiotherapy setting in patients with shoulder pain.

Yasmaine Karel
Marloes Thoomes-de Graaf
Wendy Scholten-Peeters
Paulo Ferreira
Dimitris Rizopoulos
Bart Koes
Arianne Verhagen

Published in Physiotherapy Theory and Practice
2018, 34(5)

INTRODUCTION

In physiotherapy practice patients usually follow a treatment regimen provided in coherence with the physiotherapist. This interaction between patient and therapist is referred to as a working alliance (WA). WA is first described in psychotherapy as the extent to which a client and therapist work collaboratively, purposefully and connect emotionally. WA is defined as a combination of 3 factors; agreement about the goals of treatment, the tasks of treatment and the bond between client and therapist [1].

For a treatment to be effective one important factor is that the patient complies with the regimen, after which health outcomes are more likely to improve [2]. Therefore it is essential for the therapist to provide a proper transfer of information about the goals and tasks of treatment for the patient in order to carry out the treatment regimen [3, 4]. Besides agreement about treatment goals and tasks, co-operation and compliance are achieved by means of bonding and trust between the therapist and the patient. Patients consult a physiotherapist because they seek help and they are in that case vulnerable. Help must therefore be offered and accepted based on trust. How this relationship will develop during the treatment period can have a significant impact on treatment outcome.

Several reviews have found that WA is a strong predictor of improvement in psychotherapy and psychology practices [5, 13]. Later research has established the importance of a good alliance also in other medical settings, such as in patients with ulcer disease, hypertension and diabetes [6, 7]. One review included 14 studies examining the patient-therapist relationship in physical rehabilitation setting [8]. In 9 studies a registered physiotherapist delivered the interventions. Results of the individual studies indicated that WA has a consistent positive correlation to treatment outcomes of pain, disability, physical/mental health and patient satisfaction [8]. A recent observational study of therapeutic alliance in patients with chronic low back pain confirmed these findings and found WA to be a consistent predictor of function, pain and disability measures [9]. WA might be more important in some therapies especially in those where treatment adherence represents an important component for treatment effect [10].

The Working Alliance Inventory (WAI) is one of the most commonly used and validated questionnaires to measure the working alliance [8]. It has been originally developed as a 36-item questionnaire based on Bordin's model measuring three domains; goal, task and bond [11, 12]. The WAI exists of one questionnaire for the client (WAI (C)) and one for the therapist (WAI (T)). Evidence suggests that the clients WA rating at the beginning of treatment is superior over the therapist rated version in predicting outcome [13].

The WAI was translated to Flemish, which is closely related to Dutch, named the "werk alliantie vragenlijst" (WAV). The 12 most indicative items were selected using confirmatory factor analysis to form the WAV-12 short form [14]. The WAV-12 has been used and validated in patients receiving psychotherapy in Belgium [15]. This study found a good

internal consistency for the three-factor model according to Bordin (task scale; correlation coefficient a=0.85, bond scale a=0.82, goal scale a=0.83). Correlations between the task and goal scales were good (correlation coefficient r=0.80) but correlations between the other scales were both lower (Cronbach's a=0.49). The WAV-12 used a 5-point likert scale instead of a 7-point likert scale in the original WAV-36. Therefore it is difficult to compare results from this validation study with other data. Literature does describe slightly higher correlation coefficients for the English and French short versions [14, 16]. A review has shown that translated versions of a measurement instrument for the neck do not guarantee similar measurement properties compared with the original instrument [17]. Cross-cultural validation in the Dutch population and physiotherapy setting is an important step to evaluate whether the underlying construct still holds for the WAV-12.

Therefore this study aims to investigate whether the WAV-12 is a valid measurement instrument in terms of the construct and discriminative abilities for a population of patients with shoulder pain in physiotherapy care.

METHODS

Study design

The study population consisted of patients with shoulder pain that participated in a prospective cohort study in patients consulting a physiotherapist for shoulder pain [18]. Recruitment period was from November 2011 till December 2012. The Research Committee of the Erasmus Medical Centre in Rotterdam approved the project (MEC-2011-414). After signing an informed consent patients were included and followed up for 6 months.

Participants

A total of 125 physiotherapists were invited to enrol patients. Patients consulting a physiotherapist were included if they suffered from shoulder pain, were aged ≥ 18 years and had adequate understanding of the Dutch language. Patients were excluded if they had serious pathologies (infection, cancer or fracture), surgery of the shoulder in the previous 12 months, or had received diagnostic imaging techniques such as musculoskeletal ultrasound, magnetic resonance imaging or X-ray of the shoulder in the 3 months prior to start of the study. Patients included in the cohort study were followed for 6 months and received usual physiotherapy care. Questionnaires were sent by email at 6, 12 and 26 weeks and 2 reminders were sent after 2 and 4 days whenever the patient had not responded to the questionnaire.

Working Alliance (WA)

WA was measured 6 weeks after baseline for both the patient and physiotherapist, because earlier assessment would not clearly reflect the WA. We used the Flemish version of the WAI (WAV-12). It contains 12 items scored on a 5-point scale ranging from 1 ("never") to 5 ("always") and scoring is done for the total score and each subscale (goal, task and bond). The total score ranges from 12 (low WA) to 60 (high WA), and subscales range from 4 to 20. Where the patient had to fill in the name of the therapist we replaced the empty space with the words: "my therapist".

Statistical analysis

Descriptive data for demographic and symptom severity are presented as percentages for nominal variables (gender, level of education, cause of injury, first episode, reasons for stopping treatment) and as means for continuous variables (age, symptom duration). T-tests were used to test for differences in demographics between participants scoring all WAV-12 items and those who did not. Cronbach's alpha was used to assess the internal consistency of the WAV-12 and we assessed the correlation between patient and therapist scores using Pearsons' correlation coefficient. Coefficients equal or more than 0.7 were regarded as acceptable. R and SPSS v20.0 were used to conduct the analysis.

Validation

Performance of the items in the WAV-12 questionnaire was assessed with a partial credit Rasch model [19]. The response patterns from the set of available items in the questionnaire were tested against what is expected by the model that works according to a probabilistic form of Guttman scaling [20]. This scale assumes a deterministic pattern with a hierarchical ordering of items (low and high level of item scale). When a higher level of the item is affirmed, there must be a high probability that lower items will also be affirmed. The analysis gives the probability that a person will affirm an item of the difference between the person's level of working alliance and the level of working alliance expressed by the item.

The Rasch model was used to test; 1) internal validity of the construct, 2) whether specific items exhibit different properties in different subgroups in the population (differential item functioning) and 3) whether item redundancy can be considered [21]. Analysis was done using the ltm package in the statistical programing language R [22].

Firstly a one partial credit model with the discrimination parameter fixed at one was tested to check whether it fits the data. If this model did not fit the data an extended partial credit model with a common discrimination parameter not constrained at one or separate discrimination parameters for each parameter was considered. Uni-dimensionality could further be examined to investigate if the test variance is attributable to the principal factor or construct, estimated with Cronbach's alpha. Due to the fact that some patient responses were missing, multiple imputations were utilized to calculate Cronbach's alpha.

Differential item functioning was examined based on a likelihood ratio X^2 test implemented in the lordif package in R. Expected scores for each item should remain the same whether, an older or younger person (<50, which was the mean age) and a man or women scores the same item.

Rasch analysis can be useful and psychometrically sound in modifying measurement instruments [23]. Different criteria could be considered for item redundancy: High Item Characteristic Curve (ICC), low ICC or items having similar calibrations.

RESULTS

Study population

Sixty-six physiotherapists enrolled in total 389 patients. Physiotherapists were 72% male and had a mean working experience of 15 years.

Of the 389 patients 43% were male, average age was 50 years with a mean duration of shoulder pain of 33 weeks (see table 1). At baseline only 4% of the patients did not fill out the baseline questionnaire. At 6 weeks 30% of the responses were lost to follow up.

Working alliance

Seventy-eight patients (22%) filled in all the WAV-12 questions, enabling us to calculate a total score. The mean WAV score was 45 on a total range of 24 to 60, which is slightly above 50% of the maximum score. Most patients did not answer one or more questions of the WAV-12. The population that had responded to all WAV-12 questions did not significantly differ at baseline with the patients that did not (see table 1). Even though not statistically significant, the difference for duration of complaints appeared to be large. Selective responses can therefore not be excluded. The questions with the most missing values are questions 1, 3, 7 and 9 (see figure 1). Question 3, 7 and 9 are part of the "bond" subscale and question 1 is part of the "goal" subscale. The working alliance score of therapists was 52 and for patients 45. WAV-12 scores between patient and therapist had a poor correlation (r=0.30).

Validity of WAV-12

Of all patients, 274 had at least filled in one or more items of the WAV-12. Three models were fitted to the data. The first model (RASCH) assumes the discrimination parameter is equal for all items and fixed at one. The second model (1PL) assumes the discrimination parameter is equal for all items but is estimated from the data and the third model (gpcm) assumes the discrimination parameter is free to vary across items. Likelihood ratio tests between these models showed that the third model provided the best fit to the data (p=<0.001).

Table 1. Baseline characteristics

Characteristics of cohort	Total n=389	Participants filling in all items of WAV-12; n=87	Participants, missing 1 or more items of WAV-12; n=302
Male (%)	170 (43)	41 (49)	129 (44)
Age (SD)	50 (13)	50 (14)	50 (13)
Duration of complaint in weeks (SD)	33 (82)	27 (58)	34 (88)
Comorbidity (%)			
No	128 (35)	25 (29)	103 (34)
Yes	236 (65)	62 (71)	199 (66)
Medication use (%)	183 (47)	40 (49)	144 (50)
Highest education (%)			
Primary school	40 (10)	12 (15)	28 (10)
High school	199 (51)	44 (54)	155 (54)
University or applied sciences	127 (33)	25 (31)	102 (36)
Paid job (%)	261 (67)	53 (65)	208 (72)
Profession (%)			
Physically intensive job	65 (17)	13 (25)	52 (25)
Static repetitive job	88 (23)	14 (27)	74 (35)
Job with awkward positions/postures			
Other	11 (3)	3 (6)	8 (4)
	99 (25)	22 (42)	77 (36)
NRS median (IQR)	6.0 (3.0)	6.0 (2.0)	6.0 (3.0)
SDQ (SD)	62 (23)	63 (24)	62 (23)
EQ-5D (SD)	0.83 (0.08)	0.82 (0.07)	0.83 (0.09)

NRS Numeric Rating Scale, SDQ Shoulder Disability Index, EQ-5D EuroQol 5 Dimensions, SD standard deviation

Item properties

All but two items (item 1 and 2), showed ceiling effects, meaning that most of the patients scored a good working alliance. Appendix 1 displays the item characteristic curves for the 12 items from the WAV-12. Items 5, 6 and 8 have a high slope and are endorsed at higher levels of working alliance. Items 1, 2 and 4 have a low slope (discrimination) and are endorsed at lower levels of WA. Considerable variation exists between item discrimination indicating the WAV-12 questionnaire includes items measuring the whole construct and items discriminating at lower and higher levels of working alliance (table 2). The item information curve showed the amount of information given by the questionnaire is highest between an ability of -2 and 0, implying that the item set is most useful in discriminating among individuals at the lower end of the working alliance trait.

Table 2. Discrimination values of WAV-12 items

Item	Discrimination	Standard error	Z value
1	0.496	0.103	4.793
2	0.443	0.088	5.066
3	1.286	0.225	5.716
4	0.761	0.118	6.424
5	2.212	0.457	4.842
6	2.067	0.338	6.114
7	1.377	0.234	5.895
8	2.266	0.369	6.139
9	1.151	0.208	5.537
10	1.068	0.158	6.742
11	1.414	0.224	6.319
12	1.107	0.167	6.613

Unidimensionality

Five imputed datasets were created. Cronbachs alpha's were calculated for the 12 items in each dataset and led to a pooled cronbach's alpha coefficient of 0.89. Indicating that items correlate highly and measure the same explanatory concept.

Differential Item Functioning (DIF)

The X^2 tested three models. Model 1 is a standard model where the ability for each person remains the same. Model 2 tests whether levels of ability differ among groups and model 3 adds an interaction term for the level of ability and the group in order to test whether discrimination parameters differ among groups.

Age was dichotomized in younger patients (under the mean age of 50) and older patients (50 and over). The X^2 tested flagged item one for differential item functioning where all models were statistically significant. No differential item functioning was found between men and women. Slightly higher factor scores (mean difference = 0.0385) for the WA in patients being treated by a physiotherapist with less than 13 years of experience but was not statistically significant (p=0.73).

Rasch analysis for the WAV-12 questionnaire indicates that items have good discriminative abilities for the lower end of the construct. High correlations coefficients indicate items measure one construct and other factors like age and experience of the physiotherapist did not influence item scoring. Validity for the items in the questionnaire appears to be sound but due to the difference in the percentage of missing data among the items and observed ceiling effects we advise linguistic (Dutch) and contextual (physiotherapeutic setting) adjustments.

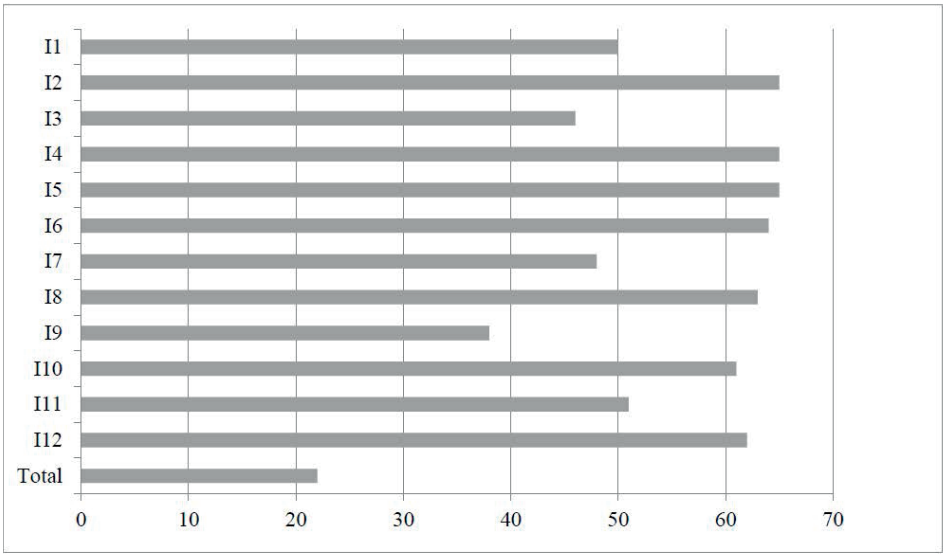


Figure 1. Relative response rate per item of WAV-12

Modification of the WAV-12

We believed rewording was necessary due to the selective number of missing responses in some items of the questionnaire and because the researchers had received comments from several patients and physiotherapists about items 3,7 and 9 of the WAV-12. Therefore we decided to make adjustments in the questionnaire and did a Delphi study. A 2 round survey was employed to ask the panels opinion on the adjustments in the WAV-12. The panel consisted of 11 members (6 clinical/research experts and 5 patients). Panel members were sent a questionnaire via email and these were sent separately to ensure panel members were unaware each of other's identity. For each item the panel member had to give his/her opinion about the adjustments with a 5-point likert scale. If the score was below 3 (neutral, disagree, totally disagree) the panel member were asked to give their reasoning and/or a suggestion for adjustment. If consensus for one item was < 80% after the first round it was included in the second round containing the suggestions of all panel members (anonymous). Full consensus (100% response rate) was reached after the second round and the adjusted questionnaire can be found in the attachments.

DISCUSSION

Main findings

Just a small proportion of patients filled in the complete WAV-12 compared to other questionnaires at 6 weeks follow-up. A large number of participants only completed a

limited number of items. This might indicate that the measurement instrument is not appropriate either in terms of language, setting, or participants had other specific reasons not to complete the questionnaire. The principal investigator also received comments from several patients and therapists, involved in the study, about items 3, 7 and 9 in the WAV-12 questionnaire. The construct theory of the WAV appeared to be sound but ceiling-effects were found in 10 items. Rewording was necessary for the WAV-12.

Comparison with the literature

Items correlated highly and measured the same explanatory concept which is found by several other translated versions of the WAI [11, 13, 14]. A French validation study found a very high correlation between the three subscales indicating that we cannot significantly distinct these subscales [16].

The poor correlation between patient and therapist WA score is consistent with other studies indicating that the two perspectives are not associated, which is confirmed by other studies as well [24, 25]. To ensure unbiased results the patient and the physiotherapist completed the rating forms independently of each other. Nevertheless, contact between the therapist and patient could not have been avoided, resulting in the possibility of deliberation between them.

WA was measured at 6 weeks when alliance might already have evolved into a stable situation whereas the first clinical experience between patient and therapist could determine more valid WA scores [26]. The literature is still inconsistent about what the optimal timing would be for measuring WA and some studies report that early WA predicted recovery after controlling for symptom change [27-30], while others have found a reduction of the predictive value of WA [31-33]. In this study WA was measured at six weeks as the first questionnaire was filled in before the first treatment. Nevertheless, we believe multiple measurements during the treatment period might yield more insight into the concept of WA.

Although WA is a valid construct within psychological interventions and research, whether it predicts recovery in a patient population in physiotherapy setting remains unknown. Psychological interventions are usually based on behavioural therapy that physiotherapists mostly use in chronic patients. The patient population in this study all have a new episode of shoulder pain where WA might be less relevant for the therapeutic process.

Strengths and limitations

This is the first study to perform a validation analysis on the Flemish version of the working alliance inventory in a physiotherapy setting. The measurement tool was able to discriminate between patients that experience a good or poor alliance. In ten items we observed ceiling effects, which might have been due to the fact that patients give socially

desirable answers or that the items do not properly assess the total construct. There appeared to be a pattern in missing items, where 4 items showed more missings than others, indicating that these might need adjustment. The questionnaire was developed in Belgium and applied in a Dutch setting which might not be appropriate given some linguistic characteristic differences of the Belgian Dutch (Flemish) and the Dutch language in the Netherlands. Due to the high number of missings in specific items (item 1, 3 and 9) and low discriminative values (item 1 and 2) we made changes in terms of adjustments in language and specific to the context of physiotherapy.

Implications for future research

The new questionnaire from our Delphi study has not been tested and therefore future research should test the psychometric properties of this questionnaire and evaluate the possible predictive value of the WA throughout the whole process of treatment in patients with musculoskeletal complaints. Whether measuring WA at the beginning or later in therapy is more predictive remains unknown. Studying a relationship between WA and recovery is complex because other factors, like self-adherence, compliance, might influence the relationship and therefore a mediation analysis might find more valid results.

CONCLUSIONS

The WAV-12 measurement tool is not suitable for implementation in clinical or research practice yet. However WA is a concept that needs attention within the field of physiotherapy and therefore we made adjustments to the questionnaire. Previous research has shown a positive correlation between working alliance and recovery in physiotherapy setting. Since shoulder pain can become a chronic condition in more than 50% of patients, interventions from physiotherapy need to be effective and a good WA can possibly contribute to optimal treatment effects.

Funding

The authors report no declarations of interest. This study is financed by the SIA-RAAK grant. The ministry of education has made this funding available for the innovation and promotion of research. This study is partly funded by a program grant of the Dutch Arthritis Foundation.

REFERENCES

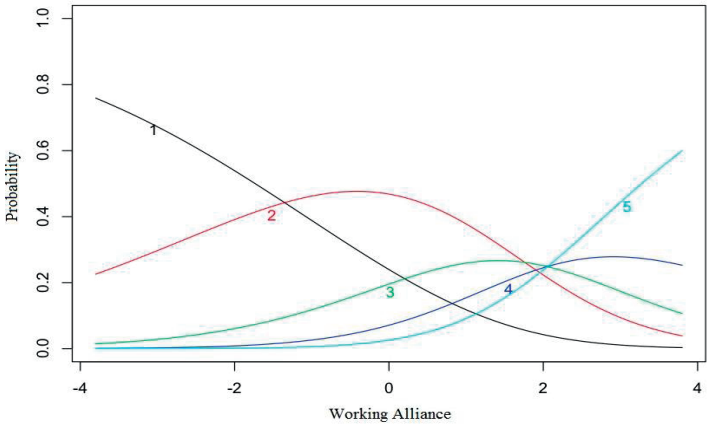
1. Bordin, E. (1979). The generalizability of the psychoanlaytic concepts of the working alliance. Psychotherapy: *Theory, Research and Practice,* 16, 252-60.
2. Bennett, J., Fuertes, J., Keitel, M., Phillips, R. (2011). The role of patient attachement and working alliance on patient adherence, satisfaction and health related quality of life in lupus treatment. *Patient education and counseling*, 85, 53-9
3. Sluijs, E., Kok, G., van der Zee, J. (1993). Correlates of Exercise Compliance in Physical Therapy. *Physical Therapy, 73*, 771-82.
4. Crow, R., Gage, H., Hampson, S., Hart, J., Kimber, A., Thomas, H. (1999). The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technology Assessment*, 3, 1-96.
5. Martin, D., Garske, J., Davis, K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 68, 438-50.
6. Kaplan, S., Greenfield, S., Ware, J. (1989). Assessing effects of physician-patient interactions on the outcomes of chronic disease. *Medical Care*. 27, S110-27.
7. Lee, Y., Lin, J. (2009). The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Social Science & Medicine*, 68, 1060-8.
8. Hall, A., Ferreira, P., Maher, C., Latimer, J., Ferreira, M. (2010). The Influence of the Therapist-Patient relationship on Treatment Outcome in Physical Rehabilitation: A Systematic Review. *Physical Therapy*, 90, 1099-110.
9. Ferreira, P., Ferreira, M., Maher, C., Refshauge, K., Latimer, J., Adams, R. (2013). The Therapeutic Alliance Between Clinicians and Patients Predicts Outcome in Chronic Low Back Pain. *Physical Therapy*, 93, 470-8.
10. Siev, J., Huppert, J., Chambless, D. (2009). The dodo bird, treatment technique, and disseminating empirically supported treatments. *The Behavior Therapist*, 32, 69-75.
11. Horvath, A., Lubrowsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology*, 61, 561-73.
12. Horvath, A., Greenberg, L. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36, 223-33.
13. Horvath, A., Symonds, D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139-49.
14. Tracey, T., Kokotovic, A. (1989). Factor strructure of the working alliance inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 207-10
15. Stinckens, N., Ulburghs, A., Claes, L. (2009). De werkalliantie als sleutelelement van de WAV-12, de Nederlandstalige verkorte versie van de Working Alliance Inventory. *Tijdschrift voor Klinische Psychologie*, 39, 44-60.
16. Corbière, M., Bisson, J., Lauzon, S., Ricard, N. (2006). Factorial validation of a French short-form of the Working Alliance Inventory. *International Journal of Methods in Psychiatric Research*, 15, 36-45.
17. Schellingerhout, J., Heymans, M., Verhagen, A., de Vet, H., Koes, B., Terwee, C. (2011). Measurement properties of translated versions of neck-specific questionnaires: a systematic review. *BMC Medical Research Methodology*, 11, 87.
18. Karel, Y., Scholten-Peeters, W., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R., van den Borne, M., Koes, B.W., Verhagen, A.P., Dinant, G.J., Tetteroo, E., Beumer, A., van Broekhoven, J.B., Heijmans, M. (2013).

Current management and prognostic facotrs in physiotherapy practice for patients with shoulder pain: design of a prospective cohort study. *BMC Musculoskeletal Disorders*, 14, 62

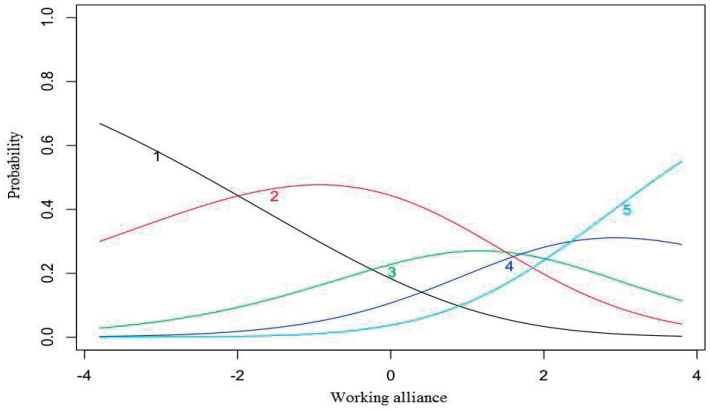
19. Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-74.
20. Guttman, L. (1950). The basis for Scalogram analysis, 4th edn. (I. S. Prediction, Ed.) New York: Wiley: The American Soldier.
21. Tennant, A., Conaghan, P. (2007) The Rasch Measurement Model in Rheumatology: What is it and When Use It? When Should It Be Applied and What Should One Look for in a Rasch Paper? *Arthritis & Rheumatism (arthritis Care & Research)*, 57, 1358-62.
22. Rizopoulos, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, 17.
23. Velozo, C., Lai, J., Mallinson, T., Hauselman, E. (2001). Maintaining instrument quality while reducing items: application of Rasch analysis to a self-report of visual function. *Journal of Outcome Measurement*, 4, 667-80.
24. Huddy, V., Reeder, C., Kontis, D., Wykes, T., Stahl, D. (2012). The Effect of Working Alliance on Adherence and Outcome in Cognitive Remediation Therapy. *The Journal of Nervous and Mental Disease*, 200, 614-9.
25. Bachelor, A. (2013). Clients' and therapists' views of the therapeutic alliance: Similarities, differences and relationship to therapy outcome. *Clinical Psychology and Psychotherapy*, 20, 118-35.
26. Horvath, A. (2001). The Alliance. *Psychotherapy*, 38,365-372.
27. Barber, J., Connolly, M., Crits-Christoph, P., Gladis, L., Siqueland, L. (2000). Alliance Predicts Patient's Outcome beyond In-Treatment Change in Symptoms. *Journal of Consulting and Clinical Psychology*, 68, 1027-32.
28. Anker, M., Owen, J., Duncan, B., Sparks, J.A. (2010). The alliance in couple therapy: Partner influence, early change, and alliance patterns in a naturalistic sample. *Journal of Consulting and Clinical Psychology*, 78, 635-645.
29. De Bolle, M., Johnson, J., De Fruyt, F. (2010). Patient and clinician preceptions of therapeutic alliance as predictors of improvement in depression. *Psychotherapy and Psychosomatics*, 79, 378-385.
30. Klein, D., Schwartz, J., Santiago, N., Vivian, D., Vocisano, C., Castonguay, L., Arnow, B., Blalock, J.A., Manber, R., Markowitz, J.C., Riso, L.P., Rothbaum, B., McCullough, J.P., Thase, M.E., Borian, F.E., Miller, I.W., Keller, M.B. (2003). Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology*, 71, 997-1006.
31. Barber, J., Luborsky, L., Crits-Christoph, P., Thase, M., Weiss, R., Frank, A., Onken, L., Gallop, R. (1999). Therapeutic alliance as a predictor of outcome in treatment of cocaine dependance. *Psychotherapy Research*, 9, 54-73.
32. DeRubeis, R., Feeley, M. (1990). Determinants of change in cognitive therpy for depression. *Cognitive Therapy and Research*, 14, 469-482.
33. Puschner, B., Wolf, M., Kraft, S. (2008). Helping alliance and outcome in psychotherapy: What predicts what in routine outpatient treatment? *Osychotherapy Research*, 18, 167-178.

Appendix 1:

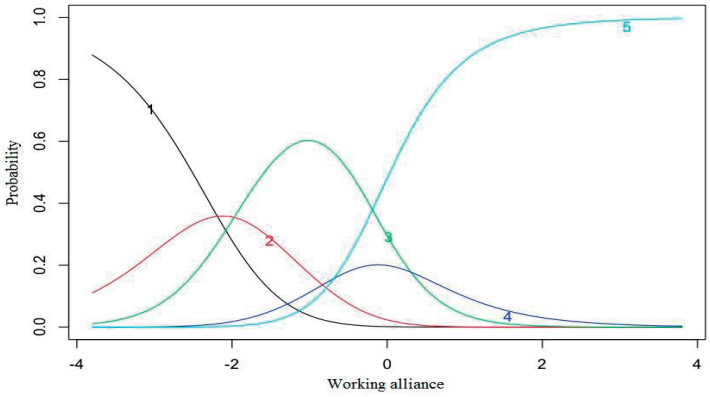
Item characteristic curves for the items in the WAV-12 questionnaire. Probability of working alliance score on the total construct for each response category of the item in different colours (1-5 likert scale).



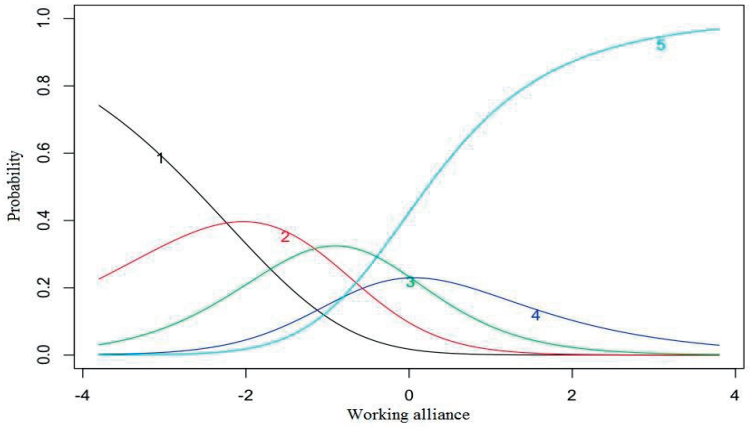
Item response category characteristic curve item 1



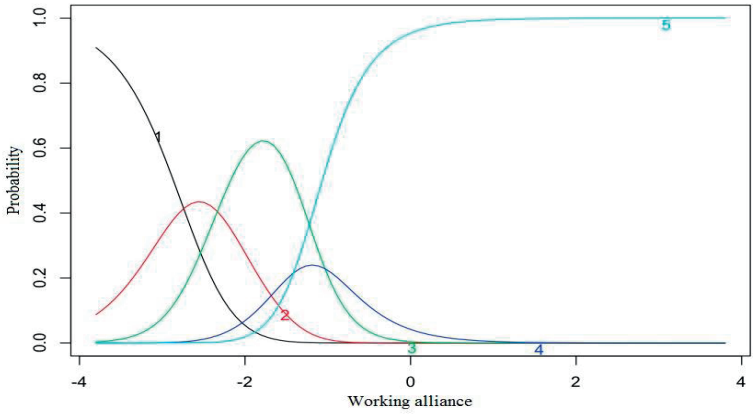
Item response category characteristic curve item 2



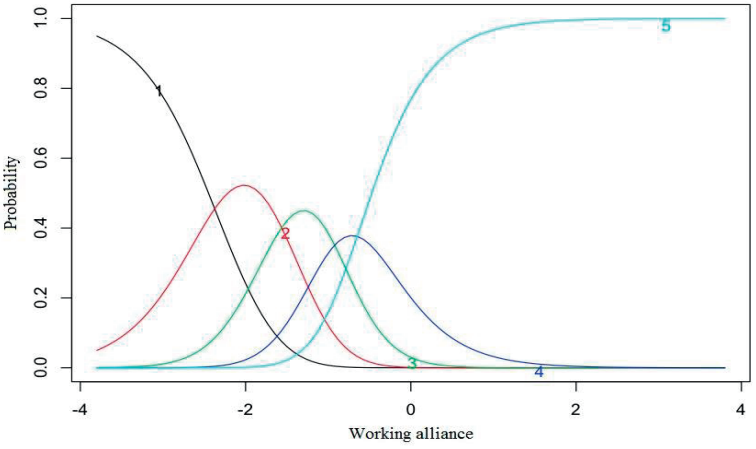
Item response category characteristic curve item 3



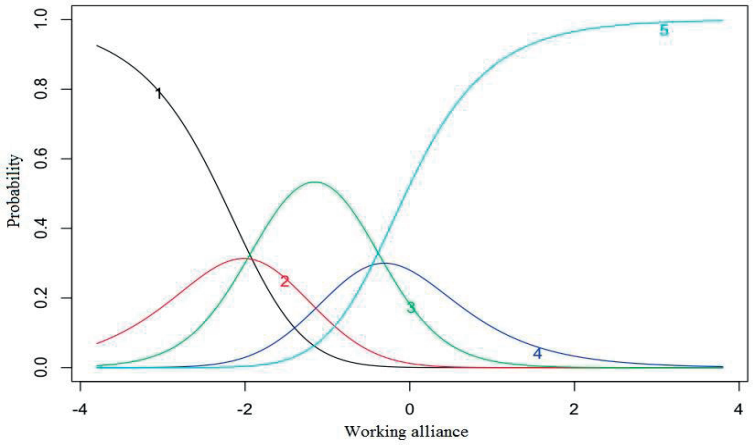
Item response category characteristic curve item 4



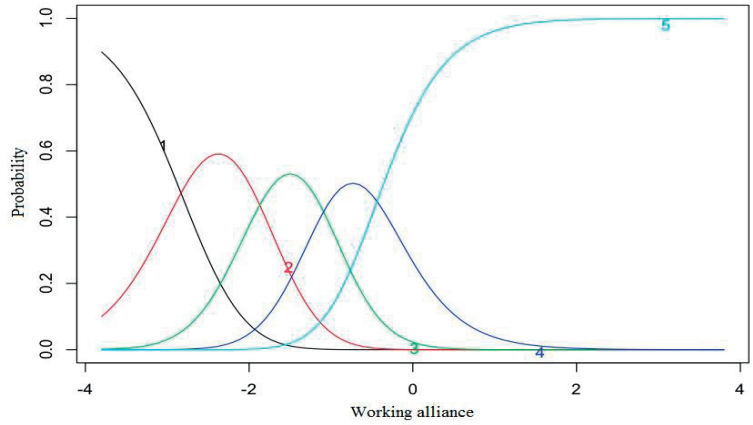
Item response category characteristic curve item 5



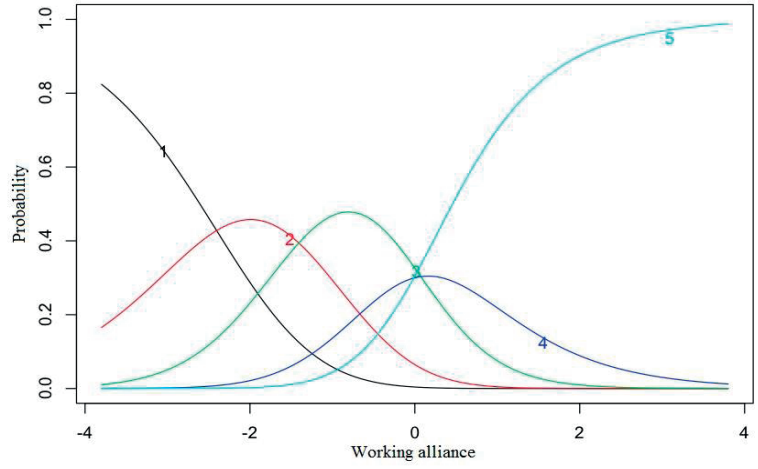
Item response category characteristic curve item 6



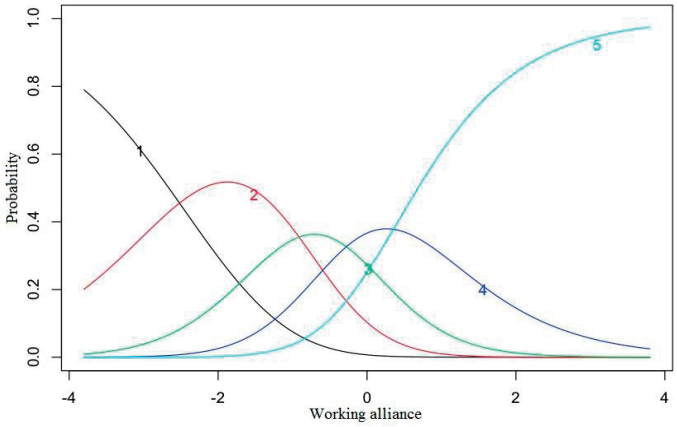
Item response category characteristic curve item 7



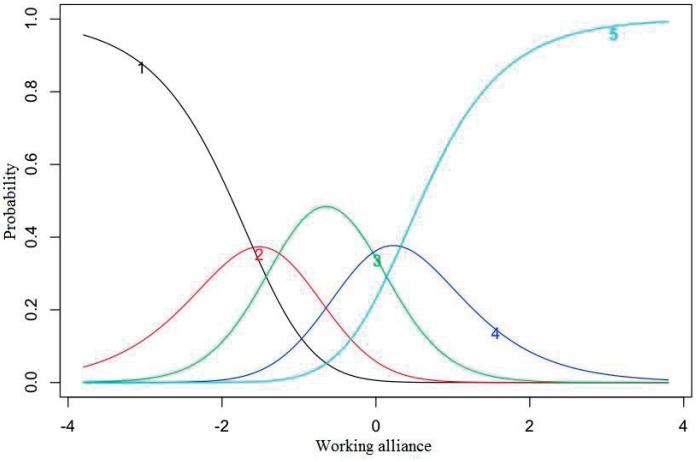
Item response category characteristic curve item 8



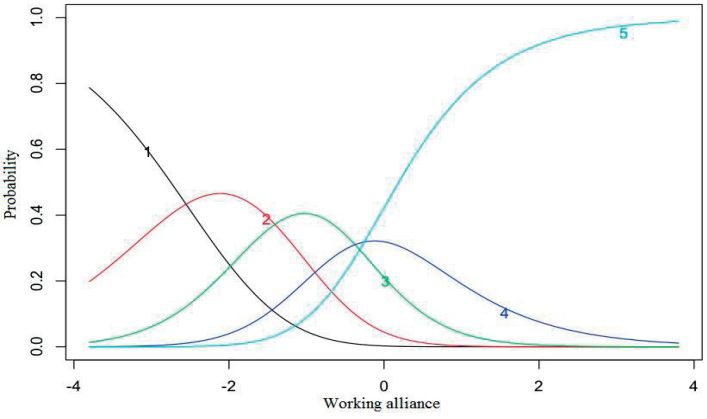
Item response category characteristic curve item 9



Item response category characteristic curve item 10



Item response category characteristic curve item 11



Item response category characteristic curve item 12