



# Cultural adaption and validation of the German version of the diagnostic thinking inventory (DTI-G)

## Ein Instrument zur Erhebung diagnostischer Kompetenz: Validierung und kulturelle Adaptation des Diagnostic Thinking Inventory (DTI-G)

Axel Georg Meender Schäfer<sup>1\*</sup>, Britta Sebelin<sup>1</sup>, Lena Spitzer<sup>2</sup>

<sup>1</sup>Hochschule für angewandte Wissenschaft und Kunst  
Hildesheim, Fakultät Soziale Arbeit und Gesundheit,  
31134 Hildesheim, Germany  
\* axel.schaefer@hawke.de

<sup>2</sup>RWTH Aachen University, Sektion Klinische Kognitionsforschung,  
Klinik für Neurologie, Medizinische Fakultät, 52074 Aachen,  
Germany

Received 10 August 2016, accepted 19 December 2018

### Abstract

Diagnostic ability is essential for expert professional practice. Several instruments have been developed to assess diagnostic skills independent of specific knowledge. One such instrument is the diagnostic thinking inventory (DTI), which is used in different settings to evaluate diagnostic performance and has shown acceptable reliability and validity.

The aim of the present study was to translate and validate a German version (DTI-G).

Cultural adaptation and translation were performed according to international guidelines. Internal consistency and item discrimination indexes were calculated. The factorial structure of the DTI-G, test-retest reliability and known-groups validity were tested.

A total of 388 physiotherapists completed the questionnaire. The internal consistency was good for the overall score of the DTI-G (Cronbach's  $\alpha = 0.84$ ). Exploratory factor analysis yielded a five-factor solution with 21 items that explained 55% of the total variance across items. A confirmatory principal component analysis resulted in the same five-factor structure, showing an acceptable to good overall fit of the model ( $CFI = 0.93$ ;  $RMSEA = 0.05$ ;  $SRMR = 0.06$ ). Test-retest reliability was found to be good (intraclass correlation coefficient  $ICC_{2,1} = 0.87$ ,  $p < 0.001$ ,  $n = 118$ ). The difference between participants with more than 9 years of clinical experience and those with less than 9 years of clinical experience (median split) was significant ( $t_{385} = 6.00$ ,  $p < 0.001$ ), supporting known-groups validity.

The results support construct validity and indicate good test-retest reliability of the DTI-G. The DTI-G can be used to measure and develop diagnostic ability of physiotherapists in clinical practice and education. Further research is necessary to validate the questionnaire for other health professions.

### Abstract

Diagnostische Kompetenz ist ein Kernelement professioneller klinischer Praxis. Verschiedene Instrumente wurden entwickelt, um diagnostische Kompetenz unabhängig von fachspezifischem Wissen zu messen. Eines davon ist das Diagnostic Thinking Inventory (DTI), das in verschiedenen Bereichen zur Überprüfung diagnostischer Kompetenz eingesetzt wird. Verschiedene Studien zeigen eine akzeptable Reliabilität und Validität des DTI. Das Ziel dieser Studie war die Übersetzung und Validierung einer deutschen Version (DTI-G).

Die Übersetzung und kulturelle Adaptation erfolgten nach internationalen Richtlinien. Interne Konsistenz, Trennschärfe der Items, Faktorstruktur, Retest-Reliabilität und known-groups validity wurden untersucht.

388 Physiotherapeuten/-innen füllten den Fragebogen aus. Es zeigte sich eine gute interne Konsistenz für die gesamte Skala (Cronbachs  $\alpha = 0.84$ ). Die exploratorische Faktorenanalyse zeigte eine fünf-faktorielle Lösung mit 21 Items und einer Varianzaufklärung von 55%. Eine konfirmatorische Hauptkomponentenanalyse bestätigte die fünf-faktorielle Struktur und zeigte eine akzeptable bis gute Modellpassung ( $CFI = .93$ ;  $RMSEA = .05$ ;  $SRMR = .06$ ). Die Retest-Reliabilität war gut mit einem  $ICC_{2,1} = 0.87$ , ( $p < 0.001$ ,  $n = 118$ ). Weiterhin zeigten sich signifikante Unterschiede zwischen Teilnehmer/-innen mit mehr und weniger als 9 Jahren Berufserfahrung ( $t_{385} = 6.00$ ,  $p < 0.001$ ). Insgesamt unterstützen die Ergebnisse die Konstruktvalidität des DTI-G und zeigen eine gute Retest-Reliabilität.

Das DTI-G kann somit die Beurteilung und Entwicklung diagnostischer Kompetenz in der physiotherapeutischen Praxis und Ausbildung unterstützen. Weitere Forschung ist nötig, um den DTI-G auch für andere Gesundheitsberufe zu validieren.

### Keywords

Diagnostic Ability – Diagnostic Thinking Inventory – Validation – Clinical Reasoning

### Keywords

Diagnostische Kompetenz – Diagnostic Thinking Inventory – Validität – Clinical Reasoning

## INTRODUCTION

Diagnostic ability is essential for expert professional practice of healthcare practitioners as it determines the course of examination and treatment and is closely related to the quality of clinical reasoning (Groves, O'Rourke & Alexander, 2003). The clinical reasoning process is highly complex and involves analytic reasoning (hypothetic-deductive reasoning), non-analytic reasoning (pattern recognition, illness scripts) and meta-cognition to monitor the process (Groves, 2012).

The question of what constitutes diagnostic expertise and how this can be measured goes back to medical research in the early 1970s (Elstein, 2009). Researchers aimed to analyse the clinical reasoning process by using simulated patients, recall tasks and verbalisation (Jensen, Gwyer, Hack & Shepard, 2007). Results showed that clinical expertise depended on quality of problem-solving or information processing. One example is Elstein's hypothetico-deductive method, consisting of cue acquisition, hypothesis generation, cue interpretation and hypothesis evaluation (Elstein, Shulman & Sprafka, 1978). Another line of thought stressed the importance of knowledge structures (Custers, Regehr & Norman, 1996), which "refers to availability of knowledge, stored in memory, during the diagnostic process" (Bordage, Grant, Marsden, 1990, p. 416). Later, Groves et al. (2003) found that the ability to integrate critical clinical information is an important sign of a competent diagnostician. Generally, he considered the following as important determinants of diagnostic ability:

- accurate collection of data
- recognition of meaningful information
- correct integration of information and
- correct interpretation of information

Diagnostic ability is commonly measured on the basis of correctly solved case examples with different formats such as script concordance tests (SCTs), clinical reasoning problems (CRPs) and key feature questions (KFQs) (Hryncak, Glover Takahashi & Nayer, 2014). However, case-based assessments are content specific, resulting in a large variability in clinicians' performance (Elstein, 2009). Case examples are closely connected to specific knowledge about a health problem, and this knowledge may be highly variable, depending on individual interest or specialisation. Whether a case example is correctly solved may depend in a higher degree on knowledge than on diagnostic ability per se.

Several instruments have since been developed to assess diagnostic skills independent of specific knowledge. One such instrument is the diagnostic thinking inventory (DTI) (Bordage, Grant & Marsden, 1990). The DTI consists of 41 items representing two dimensions "flexibility in thinking" and "evidence for structure in

memory". Flexibility in thinking describes the skill to change between deterministic and responsive modes of information acquisition. In the deterministic mode, clinicians ask questions based on memorised knowledge (Bordage et al., 1990). In the responsive mode, clinicians are able to set aside the deterministic mode and immediately react on new information that may alter the direction of reasoning.

The DTI was used in different settings to evaluate diagnostic performance of medical students and practitioners (Beullens, Struyf & Van Damme, 2006; Gehlhar, Klimke-Jung, Stosch & Fischer, 2014; Goss, Reid, Dodds & McColl, 2011; Lee et al., 2010; Rahayu & McAleer, 2008). The reliability and validity of the DTI were also shown for physiotherapists (Jones, 1997) with satisfactory test-retest reliability and significant differences between students and qualified physiotherapists.

The aims of the present study were to translate the DTI to German and to evaluate its validity for allied health professionals in Germany.

The following objectives were pursued:

1. to create a culturally adapted German translation of the DTI following the recommendations of Beaton, Bombardier, Guillemin and Ferraz (2000)
2. to adapt the DTI-G according to analysis of internal consistency and factorial structure and
3. to investigate construct validity and test-retest reliability of the German translation.

## METHODS

### Translation and cultural adaptation

The adaption and translation of the DTI were performed according to Beaton et al. (2000), involving two forward translations in German language, synthesis of the translations, two back translations in English and consolidation of the German versions involving the original developer. Bilingual translators whose mother language is German produced the two forward translations. One of them is a professional translator (university diploma), and the other is a speech and language therapist. The pre-final version of the DTI-G was then pretested in the target group. The pretest focussed on relevance, comprehensibility and identification of unclear formulations or questions that are irrelevant for the target group. Twelve cognitive interviews were conducted using the think-aloud method (Willis, 2005) with four occupational therapists, four physiotherapists and four speech and language therapists. The pre-test sample consisted of two males and ten females with work experience ranging from 0 to 21 years. The levels of education varied, including vocational education, bachelor's degree and master's degree. Written informed



consent was obtained from each participant. The subjects were asked to complete the DTI-G and to think aloud while responding (Willis, 2005). The verbal reports were audiotaped, transcribed and analysed according to the methods of qualitative content analysis (Mayring, 2010). Following the analysis of the pretest data, the wording of a number of items was adapted (see the Results section).

### Data collection and procedure

After creating an online version of the DTI-G with the open-source program LimeSurvey (LimeSurvey GmbH, Hamburg, Germany), physiotherapists and speech and language therapists were contacted by e-mail and asked to complete the DTI-G. Additionally, the link to the online questionnaire was published on the homepages of professional associations. Four universities were contacted, which forwarded the link on to their students. After completing the questionnaire, participants were asked if they would be willing to participate in the retest study after four weeks.

### Construct validity

An exploratory factor analysis (EFA) using the principle components extraction and orthogonal varimax rotation was performed on the 41 items to explore the instrument's factorial structure (Bühner, 2011). We decided to use principal components analysis based on the assumption that items are measured without error. In this case, principal components analysis provides the most parsimonious description of the data. Factors were extracted if their eigenvalue was  $>1$ . Oblique rotations were used when assumptions with regard to the factorial structure were not met. Items were deleted if their factor loading was below 0.30.

To examine the discriminatory power of each single question/item, the item discrimination index (IDI) was calculated. This index indicates the extent to which a single item measures the construct. Items were deleted if the IDI was below 0.20 (Bühner, 2011). Furthermore, face validity of questionable items was assessed within the research team based on logical and professional considerations.

### Reliability

Internal consistency of the DTI-G was assessed by calculating Cronbach's alpha, alpha if item deleted, item-scale correlations, and item difficulty. Cronbach's alpha indicates if individual items measure the same concept (Bühner, 2011).

To assess four-week test-retest reliability, the two-way random intraclass-correlation coefficient ( $ICC_{2,1}$ ) was

calculated for the total score of the DTI-G. A two-way random model was chosen because raters were chosen randomly and answered each of the items (de Vet et al., 2011).

### Known-groups validity

Additionally, known-groups validity was tested based on the duration of clinical experience. Subjects were divided into two groups using the median split method. Group differences were tested using the independent t-test.

### Data analysis

After completion of data collection, data were analysed using the Statistical Package for the Social Sciences (SPSS version 21; IBM Inc., New York, USA). The two-sided significance level was set at  $\alpha = 0.05$ .

## RESULTS

### Comprehensibility and relevance (pretest)

The twelve interviews with physiotherapists, speech and language therapists and occupational therapists showed differences in the diagnostic process and thinking between the professionals. All four occupational therapists had difficulties to answer questions about diagnosis, differential diagnosis or procedure. They explained that occupational therapists basically have a different diagnostic process compared to other health professionals and that the focus is rather on activity than on symptoms and pathology. This finding indicates that two different versions of the questionnaire are needed: one version for speech and language therapists and physiotherapists and another version for occupational therapists; the latter was not within scope of this study. Therefore, occupational therapists were not included in the further validation process.

Generally, analyses of the eight interviews with physiotherapists and speech and language therapists showed that the majority of the questions were understandable. Ten of the 41 items had to be revised because the wording was not comprehensible. Questions were reworded to reflect a more practice-oriented language (see Appendix 1 for the translation of the DTI-G).

### Study sample

The link to the online version of the DTI-G was sent to approximately 10,000 speech and language therapists and 30,000 physiotherapists in Germany. In all, 800 ( $>2\%$ ) questionnaires were returned. A sample consisting of n

= 388 physiotherapists (58% females) and n=122 speech and language therapists (88% females) completed the whole questionnaire. Participating physiotherapists had a mean age of 37.23 years ( $SD = 11.07$ ) and a mean clinical experience of 12.91 years ( $SD = 9.87$ ). Speech and language therapists had a mean age of 31.53 years ( $SD = 11.36$ ) and a mean clinical experience of 6.06 years ( $SD = 8.34$ ). After four weeks, 118 physiotherapists and 54 speech and languages therapists completed the questionnaire again. As the number of questionnaires returned by speech and language therapists was too low, only questionnaires of physiotherapists were included in the EFA. According to Bühner (2011), a good sample size for EFA should exceed 300.

### Construct validity

#### Item discrimination index

IDI for the original 41 items ranged from 0.04 to 0.62 for the subscale “structure in memory” and from -0.10 to 0.57 for the subscale “flexibility in thinking” (table 1). Items Q3, Q11, Q14, Q24 and Q39 were removed as their IDI was below 0.20 (Bühner, 2011), and consensus was reached within the research team that content validity of these items was questionable. Items Q17 and Q36 were retained although their IDI was below 0.20 as they were considered crucial for the underlying construct.

### Exploratory Factor Analysis (EFA)

A principle component factor analysis (PCA) was used to determine the underlying constructs of the scale. The Kaiser–Meyer–Olkin measure of sampling adequacy was 0.89, which suggests that the data were suitable for factor analysis. Bartlett’s test of sphericity reached statistical significance ( $p < 0.001$ ), which supported the factor ability of the correlation matrix.

A PCA with promax oblique rotation was carried out for 36 items. Oblique rotations are recommended when assumptions in regard to the factorial structure are not made (Bühner, 2011). The PCA resulted in a five-factor model. A factor loading of 0.30 was used as the cutoff score for selecting the items (Bühner, 2011). Thus, items in this study that had a factor loading of 0.30 or higher were retained (Table 2). All other items, except items 6 and 16, were removed, resulting in remaining 21 items. Items 6 and 16 were considered substantial to the overall construct and were retained, although factor loadings were  $> 0.20$  and  $< 0.30$ .

A PCA with promax oblique rotation with the remaining 21 items resulted in a five-factor solution that explained 55% of the total variance across all items. Factor 1 accounted for 29% of the total variance, factor 2 accounted for 9% of the total variance, factor 3 accounted

*Table 1: Item-scale statistics (see Appendix 1 for the translation of the items)*

Item	Item Discrimination Index	Cronbach's alpha when item deleted
<i>Original Factor 1: knowledge structure in memory</i>		
Q1	0.32	0.77
Q7	0.25	0.77
Q8	0.46	0.75
Q9	0.50	0.75
Q10	0.49	0.76
Q12	0.55	0.75
Q13	0.60	0.75
Q14	0.11	0.78
Q17	0.04	0.78
Q18	0.34	0.76
Q19	0.30	0.77
Q20	0.26	0.77
Q21	0.39	0.76
Q22	0.29	0.77
Q25	0.37	0.76
Q29	0.48	0.76
Q31	0.62	0.75
Q33	0.57	0.75
Q37	0.21	0.78
Q39	0.12	0.78
<i>Original Factor 2: flexibility in thinking</i>		
Q2	0.21	0.73
Q3	-0.09	0.77
Q4	0.36	0.72
Q5	0.54	0.71
Q6	0.37	0.72
Q11	-0.10	0.76
Q15	0.48	0.71
Q16	0.22	0.73
Q23	0.21	0.73
Q24	0.10	0.74
Q26	0.56	0.71
Q27	0.50	0.71
Q28	0.24	0.73
Q30	0.50	0.71
Q32	0.33	0.72
Q34	0.26	0.73
Q35	0.28	0.73
Q36	0.18	0.73
Q38	0.54	0.71
Q40	0.45	0.72
Q41	0.57	0.70



Table 2: Results of the exploratory factor analysis

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Name of factor
Q5	0.43			0.23	0.24	Recognition of meaningful information
Q10	0.82					
Q12	0.39					
Q13	0.72					
Q15	0.62					
Q33	0.38					
Q4		0.40				Information processing
Q22		0.35				
Q26	0.21	0.71				
Q29		0.62				
Q30		0.73				
Q37	-0.24	0.39		0.26		
Q6		0.25	0.22			Flexibility in thinking
Q16			0.27			
Q34			0.62			
Q38		0.21	0.51			
Q1				0.75		Structure in knowledge
Q7				0.65		
Q20				0.68		
Q8					0.66	Memory retention
Q9					0.61	
Eigenvalue	6.06	1.99	1.03	1.30	1.16	
% total variance	28.69	9.49	4.90	6.18	5.54	

for 5% of the total variance, factor 4 accounted for 6% of the total variance and factor 5 accounted for 6% of the total variance (table 2).

### Labels of the new factors

Factors were given names that reflect the meaning of the respective items (see Appendix 2 for the translated items of the final version of the DTI-G). The labels of factors 1 and 2 were selected based on findings of Groves et al. (2003), and labels of factor 3 and 4 were based on those used by Bordage et al. (1990). Factor 5 was labelled based on consensus within the research team. Factor 1 was labelled “Recognition of meaningful information”; factor 2 “Information processing”, factor 3 “Flexibility in thinking”, factor 4 “Structure in knowledge” and factor 5 “Memory retention”.

### Confirmatory factor analysis (CFA)

A confirmatory principal component analysis with oblique rotation was carried out with the retest sample (n=176

consisting of 118 physiotherapists and 58 speech and language therapists). The maximum likelihood factor analysis with oblique rotation was used to test the goodness of fit of the five-factor model. The results, including the loading of each item on the factors, the correlation between the five factors, and the CFA model fit parameters, are presented in Figure 1. The results show an acceptable to good overall fit ( $CFI = 0.93$ ;  $RMSEA = 0.05$ ;  $SRMR = 0.06$ ) of the model (Figure 1) (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger & Müller, 2003).

### Internal consistency

The internal consistency of the DTI-G 5-factor model with 21 items was considered as good with Cronbach’s  $\alpha = 0.84$ .

### Test-retest reliability

The four-week test-retest reliability was found to be good for the 21-item DTI-G with an  $ICC_{2,1} = 0.87$  (95% CI 0.81–0.91),  $p < 0.001$  and  $n = 118$  physiotherapists.

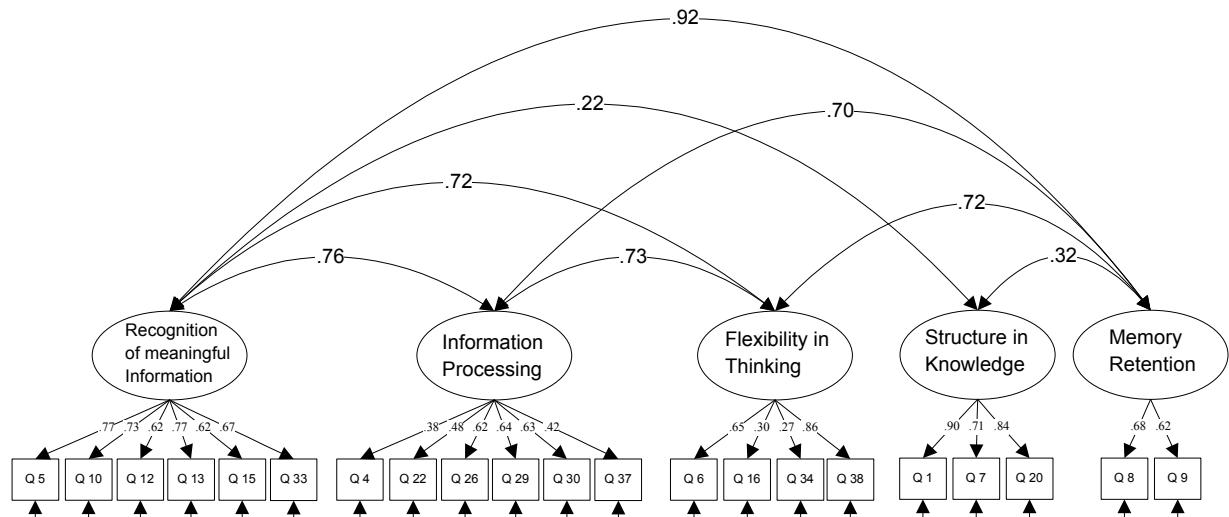


Figure 1: CFA of 5-factor model with retest sample.  $n = 176$ ,  $CFI = 0.93$ ,  $RMSEA = 0.05$  and  $SRMR = 0.06$ .

## Known-groups validity

Additionally, known-groups validity of the 21-item DTI-G was tested. Following median split, the difference between participants with equal or more than 9 years of clinical experience ( $DTI = 97.90$ ,  $SD = 13.83$ ) and those with less than 9 years of clinical experience ( $DTI = 88.55$ ,  $SD = 16.81$ ) was significant ( $t_{385} = 6.00$ ,  $p < 0.001$ , 95% CI 7.82–10.88).

## DISCUSSION

The aims of the study were to translate the DTI (Bordage et al., 1990) into German and to analyse its psychometric characteristics for use with German physiotherapists. After translation into German, the DTI-G was pretested in the target group using think-aloud interviews. The final online version of the DTI-G was completed by 388 physiotherapists. The sample size and composition were heterogeneous regarding training qualification and work experience.

EFA of the German translation of the DTI did not reproduce the original two-factor structure with 41 items (Bordage et al., 1990). A total of 20 items were removed due to low IDI ( $< 0.20$ ) or low factor loading ( $< 0.30$ ), resulting in a final five-factor model with 21 items. This model demonstrated good internal consistency, an acceptable to good overall fit, good test-retest reliability and good known-groups validity.

### Previous research into validity of the DTI

The internal consistency was good for the overall score (Cronbach's  $\alpha = 0.84$ ) corresponding with the findings

from other studies (Bordage et al. 1990, Jones 1997, Gelhar et al. 2014).

Another aim of our study was to investigate factorial structure of the DTI. However, the two-factor structure (*flexibility in thinking* and *structure in memory*) of the original DTI could not be confirmed. Our findings indicate a five-factor structure. To date, studies investigating the factorial structure of the DTI could not be detected; thus, comparison to the results of other studies is not possible.

### Strengths and limitations

The process of the translation and cross-cultural adaption was performed according to international recommendations (Beaton et al. 2000). The back translation was authorised by Bordage, the original author of the DTI. This approach guarantees that content validity of the inventory was maintained within the translation process. The cognitive interviewing technique, which was used in the pretest of our study, additionally supports content validity of the DTI-G.

The DTI-G consists of 21 items, which might be more practical and time-efficient for the use in future research projects than the original version with 41 items. However, removing 20 items from the original questionnaire might raise the question if the German version still measures the same construct. Two points need to be taken in account in regard to this question. First, to our knowledge, IDIs and factorial structure of the original English DTI have never been analysed. Doing so could also result in reduction of items. Second, many questionnaires have short forms that have been shown to measure the same construct as the long version. To strengthen construct validity, future studies should investigate concurrent validity of the



DTI-G by comparing it to other instruments known to measure diagnostic ability.

Furthermore, the retest sample completed the 41-item DTI-G after four weeks. As we calculated the ICC for the 21-item version, this measure could be biased. Another reliability study with the 21-item instrument should be carried out.

### **Application of the DTI-G**

The DTI-G is a reliable and valid inventory for use with German physiotherapists. It can be used to evaluate teaching concepts and improve diagnostic thinking in physiotherapy. This study focused on the validation of the DTI-G only for physiotherapists. Thus, the instrument should not be used for research in other professions.

### **Recommendations for future research**

Further research is necessary to evaluate convergent and divergent validity and validate the DTI-G for other health professions. In addition, sensitivity to change in the DTI-G needs to be explored. Further studies will be necessary to

validate the DTI-G for other health professionals such as speech and language therapists, occupational therapists and physicians.

In addition, future studies should be conducted to generate norm data, in order to interpret the results of individual subjects.

### **CONCLUSION**

The DTI-G can be used to measure and develop diagnostic ability of physiotherapists in clinical practice and education. Further research is necessary to evaluate convergent and divergent validity and validate the DTI-G for other health professions.

### **ACKNOWLEDGEMENTS**

The authors would like to thank Carsten John, Kathrin Allgöwer, Marion Grafe, Katharina Matzel and Katharina Maybaum who gave their extensive personal and professional guidance in the adaption and translation process.

### **References**

- Beaton, D.E., Bombardier, C., Guillemin, F. & Ferraz, M.B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*, 25(24), 3186-91.
- Beullens, J., Struyf, E. & Van Damme, B. (2006). Diagnostic ability in relation to clinical seminars and extended-matching questions examinations. *Medical Education*, 40(12), 1173-9.
- Bordage, G., Grant, J. & Marsden, P. (1990). Quantitative Assessment of Diagnostic Ability. *Medical Education*, 24, 413-25.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. 3rd ed. München: Pearson.
- Custers, E.J.F.M., Regehr, G. & Norman, G.R. (1996). Mental representations of medical diagnostic knowledge: a review. *Acad Med*, 71, 55-61.
- de Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). Reliability. In *Measurement in Medicine* (pp. 96-148). Cambridge: Cambridge University Press.
- Elstein, A.S. (2009). Thinking about diagnostic thinking: a 30-year perspective. *Adv Health Sci Educ Theory Pract.*, 14 Suppl 1, 7-18. doi: 10.1007/s10459-009-9184-0
- Elstein, A. S., Shulman, L.S. & Sprafka, S.A. (1978). *Medical Problem Solving: An Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press.
- Gehlhar, K., Klimke-Jung, K., Stosch, C. & Fischer, M.R. (2014). Do different medical curricula influence self-assessed clinical thinking of students? *GMS Z Med Ausbild*, 31(2), 1 - 20. doi: 10.3205/zma000915
- Goss, B., Reid, K., Dodds, A. & McColl, G. (2011). Comparison of medical students' diagnostic reasoning skills in a traditional and a problem based learning curriculum. *International Journal of Medical Education*, 2, 87-93. doi: 10.5116/ijme.4e64.055a
- Groves, M. (2012). Understanding clinical reasoning: the next step in working out how it really works. *Medical Education*, 46(5), 444-6. doi: 10.1111/j.1365-2923.2012.04244.x
- Groves, M., O'Rourke, P. & Alexander, H. (2003). The clinical reasoning characteristics of diagnostic experts. *Med Teach*, 25(3), 308-13. doi: 10.1080/0142159031000100427
- Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. doi:10.1080/10705519909540118
- Hryncak, P., Glover Takahashi, S. & Nayer, M. (2014). Key-feature questions for assessment of clinical reasoning: a literature review. *Medical Education*, 48(9), 870-83. doi: 10.1111/medu.12509.
- Jensen, G.M., Gwyer, J., Hack, L.M. & Shepard, K. F. (2007). *Understanding Expertise: Connecting Research and Theory to Physical Therapy*. In: W.B. Saunders (Ed.), *Expertise in Physical Therapy Practice* (pp. 19-47). 2nd ed. St. Louis: Elsevier.
- Jones, U. F. (1997). The reliability and validity of the Bordage, Grant & Marsden diagnostic thinking inventory for use with physiotherapists. *Med Teach*. 19(2), 133-40.
- Lee, A., Joynt, G.M., Lee, A.K.H., Ho, A.M.H., Groves, M., Vlantis, A.C., et al. (2010). Using illness scripts to teach clinical reasoning skills to medical students. *Family medicine*, 42(4), 255-61.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 11th ed. Weinheim, Basel: Beltz.

- Rahayu, G.R. & McAleer, S. (2008). Clinical reasoning of Indonesian medical students as measured by diagnostic thinking inventory. *South East Asian Journal of Medical Education*, 2(1), 42-7.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23-74.
- Willis, G. B. (2005). Cognitive Interviewing. A Tool for Improving Questionnaire Design: Sage: Thousand Oaks.

## APPENDIX 1

### **Diagnostic Thinking Inventory (DTI-G): full version in German**

Anmerkung: Der in diesem Fragebogen verwendete Begriff Diagnose bezieht sich auf die logopädische bzw. physiotherapeutische Diagnose bzw. den Befund.

#### **1. Wenn der Patient seine Symptome<sup>1</sup> angibt,**

denke ich an die Symptome im Wortlaut  
des Patienten.

denke ich an die Symptome in  
abstrakteren Ausdrücken (z. B.: „akut“  
anstatt „seit 4 Tagen“, „bilateral“ anstatt  
„beide Hände“).

#### **2. Wenn ich mögliche Diagnosen abwäge,**

versuche ich diese nach ihrer Wichtigkeit  
zu ordnen.

versuche ich jeder einzelnen gleiche  
Wichtigkeit oder Gewichtung zu geben.

#### **3. Beim Überlegen möglicher Diagnosen**

denke ich beim jeweiligen Fall bereits sehr  
früh an mögliche Diagnosen.

sammle ich zuerst alle klinischen Daten  
und denke dann darüber nach.

#### **4. Beim Anamnesegespräch**

bleibe ich oft bei einer möglichen  
Diagnose.

fällt es mir leicht, unterschiedliche  
Diagnoseansätze in Betracht zu ziehen.

#### **5. Während des Anamnesegesprächs,**

wenn ich dem Gedankengang des  
Patienten folge, verliere ich gelegentlich  
meinen roten Faden.

kann ich meine Ideen weiterverfolgen,  
selbst wenn ich dem Gedankengang des  
Patienten folge.

1 Zur besseren Lesbarkeit wird im Folgenden das männliche Geschlecht verwendet. Selbstverständlich sind trotz der gewählten Form immer beide Geschlechter gemeint



**6. Wenn ich mich für eine Diagnose entscheiden soll,**

stört es mich nicht, meine Entscheidung auf später zu verschieben.

habe ich das Gefühl, eine Entscheidung treffen zu müssen, auch wenn ich mir nicht sicher bin.

**7. Wenn der Patient seine Zeichen und Symptome geschildert hat,**

denke ich in seinen/ihren Worten darüber nach.

übersetze ich diese in medizinische Fachausdrücke (z.B. „Taubheitsgefühl“ wird zu „Parästhesie“).

**8. In Bezug auf die allgemeine Krankengeschichte**

habe ich das Gefühl, dass ich diese oft nicht ausreichend aufnehme.

nehme ich diese zu meiner eigenen Zufriedenheit meistens ausreichend auf.

**9. Während der Patient seine gesamte Krankheitsgeschichte erzählt,**

fällt es mir oft schwer, mich zu erinnern, was gesagt wurde.

schaffe ich es normalerweise, mir zu merken, was gesagt wurde.

**10. Im Verlauf des Patientengesprächs stelle ich fest,**

dass mir einige Schlüsselinformationen ins Auge zu springen scheinen.

dass es oft schwierig ist, zu wissen, auf welche Informationen ich mich konzentrieren sollte.

**11. Wenn die vom Patienten beschriebenen Symptome für mich keinen Sinn ergeben,**

dann suche ich weiter nach neuen Informationen, um auf neue Ideen zu kommen.

bitte ich den Patienten, die genannten Symptome deutlicher zu beschreiben.

**12. In der Abwägung möglicher Diagnosen**

ziehe ich oft nicht zutreffende Diagnosen in Betracht.

bewegen sich meine Diagnosen häufig in die richtige Richtung.

**13. Während ich Informationen über den Patienten sammle,**

ordnen sich diese meist von selbst in meinen Gedanken.

fällt es mir schwer, Zusammenhänge zwischen den einzelnen Informationen zu erkennen.

**14. Wenn sich eine Diagnose herausstellt, an die ich ursprünglich nicht gedacht hatte,**

dann war mir diese Erkrankung zwar bekannt, nur habe ich sie nicht in Betracht gezogen.

dann habe ich meist nicht genug über diese Erkrankung gewusst.

**15. Während eines Anamnesegesprächs**

kann ich irrelevante Informationen nicht ausblenden.

fällt es mir leicht zu entscheiden, welche Information ich weglassen kann, da sie irrelevant sind.



---

**16. Wenn die vom Patienten geschilderten Zeichen und Symptome keinen Sinn ergeben,**

dann fahre ich fort, um neue Informationen zu erhalten und eine neue Perspektive einzunehmen.

dann betrachte ich die genannten Symptome aus einer anderen Perspektive, bevor ich fortfahre weitere Informationen zu sammeln.

**17. Wenn ich eine Reihe von möglichen Diagnosen in Betracht ziehe,**

stehen diese eher miteinander im Zusammenhang.

gehen diese eher in verschiedene Richtungen.

**18. Wenn mir eine mögliche Diagnose in den Sinn kommt,**

erwarte ich meist mögliche Zeichen und Symptome, die mit der Diagnose einhergehen.

hilft mir dies häufig nicht zu entscheiden, was ich den Patienten als Nächstes fragen möchte.

**19. Wenn ich nur sehr wenig über eine bestimmte Art von Krankheit weiß,**

ist es mir dennoch möglich, zu einer Diagnose zu gelangen.

habe ich große Schwierigkeiten, zu einer Diagnose zu gelangen.

**20. Unter Berücksichtigung der Zeichen und Symptome des Patienten**

denke ich über diese im genauen Wortlaut des Patienten nach.

denke ich an diese in möglichen Gegensätzen (z. B.: spastisch - schlaff, fortschreitend – plötzlich, unilateral - bilateral).

**21. Wenn ich viel über eine bestimmte Art von Krankheit weiß und eine Diagnose stellen soll,**

dann fällt es mir leicht, zu einer Diagnose zu kommen.

dann bin ich eher unstrukturiert und es fällt mir schwer, mich auf eine Diagnose festzulegen.

**22. Wenn ich im Verlauf der Anamnese bereits einige Ideen über mögliche Diagnosen habe,**

geben mir neue Informationen zusätzliche Ideen.

geben mir neue Informationen keine zusätzlichen Ideen.

**23. Wenn ich eine Anamnese erhebe,**

dann kommen mir neue Ideen beim Durchgehen der bestehenden Informationen.

benötige ich neue Informationen, um neue Ideen für den Fall zu entwickeln.

**24. Wenn sich der Patient bei der Schilderung der Symptome ungenau bzw. mehrdeutig ausdrückt,**

dann lasse ich ihn weitererzählen.

unterbreche ich und bitte den Patienten, mir diesen Sachverhalt noch einmal genau zu erklären.

**25. Nach einem Anamnesegespräch**

fallen mir selten andere Dinge ein, die ich in Bezug auf die Krankheit des Patienten hätte fragen sollen.

fallen mir häufig andere Dinge ein, die ich in Bezug auf die Krankheit des Patienten hätte fragen sollen.



**26. Wenn ich eine Information erhalte, die mich an eine mögliche Diagnose denken lässt,**

denke ich an frühere Informationen zurück, um abzuklären, ob diese damit zusammenpassen oder nicht.

überdenke ich selten die Informationen, die ich vorher gesammelt habe.

**27. In Bezug auf die Diagnose, die ich schlussendlich stelle,**

habe ich selten Zweifel.

fühle ich mich häufig zu unsicher.

**28. Beim Treffen einer diagnostischen Entscheidung**

entscheide ich, indem ich jede mögliche Diagnose für sich selbst betrachte.

entscheide ich, indem ich mögliche Diagnosen vergleiche und gegenüberstelle.

**29. Wenn ich viel über eine bestimmte Krankheit weiß und eine Diagnose stellen muss,**

überprüfe ich die meisten Möglichkeiten, bevor ich eine Entscheidung treffe.

habe ich oft viele unterschiedliche Ideen, die ich aber nicht weiter verfolge.

**30. Während mir der Fall klarer wird,**

halte ich es nicht für hilfreich, während des Prozesses die Informationen zusammenzufassen.

mache ich regelmäßig Bestandsaufnahmen der Daten und meiner Ideen.

**31. Wenn ich zu einer Diagnoseentscheidung komme,**

gibt es oft bereits erhobene Informationen, die ich einfach wieder vergessen habe.

habe ich normalerweise alle Informationen in Betracht gezogen.

**32. Wenn ich eine Idee habe, was dem Patienten fehlen könnte,**

fühle ich mich am wohlsten, wenn ich diese Idee ohne Ablenkung weiterverfolgen kann.

fühle ich mich wohl, wenn ich auch andere Optionen in Betracht ziehe, um dann später zu meiner Ursprungsidee zurückzukommen.

**33. Wenn ich eine allgemeine Vorstellung habe, was dem Patienten fehlen könnte,**

kann ich gewöhnlich fortfahren, bis ich eine spezifische Diagnose gefunden habe.

fällt es mir schwer, eine spezifische Diagnose zu formulieren.

**34. Während des Patientengesprächs**

schaffe ich es, meine Ideen zu überprüfen, selbst wenn ich dem Patienten die Gesprächsführung überlasse.

bin ich nur dann erfolgreich, wenn ich die Richtung des Gesprächs vorgeben kann.

**35. Wenn ich aus verschiedenen Möglichkeiten eine Diagnose auswähle,**

bin ich für gewöhnlich nicht in der Lage, irgendeine meiner Ideen gänzlich auszuschließen.

bin ich in der Lage, die meisten meiner Ideen komplett auszuschließen.

**36. Wenn meine Entscheidung über einen Patienten feststeht,**

wäre ich bereit, meine Entscheidung zu ändern.

würde ich meine Entscheidung ungern ändern wollen.

### **37. Wenn ich meine diagnostischen Ideen in Betracht ziehe, mache ich das auf der Basis**

der gesamten Informationen bis zu diesem Zeitpunkt.

einiger herausstechender Zeichen und Symptome.

### **38. Wenn ich nicht weiß, welche Schlüsse ich aus dem Gespräch ziehen soll,**

fällt es mir leicht, die Informationen auf eine neue Weise zu betrachten.

fällt es mir schwer, die Informationen auf eine neue Weise zu betrachten.

### **39. Wenn ich medizinische Befunde (z. B. MRT, CT, Laborbefunde) anfordere,**

mache ich dieses als Teil der klinischen Routineuntersuchung.

mache ich dieses, um spezifische Informationen und unterstützende Belege zu erhalten.

### **40. Beim Erwägen möglicher Diagnosen**

vergleiche ich die möglichen Diagnosen und stelle sie gegenüber.

betrachte ich jede Diagnose für sich selbst.

### **41. Im Hinblick auf die Art, wie ich ein Patientengespräch führe,**

sammle ich für gewöhnlich während des Gespräches alle nötigen Informationen.

stelle ich häufig nicht alle Fragen, die ich zu der Zeit stellen sollte.

## **APPENDIX 2**

### **Diagnostic Thinking Inventory (DTI-G): final version in German**

#### **1. (1)<sup>1</sup> Wenn der Patient<sup>2</sup> seine Symptome angibt,**

denke ich an die Symptome im Wortlaut des Patienten.

denke ich an die Symptome in abstrakteren Ausdrücken (z. B.: „akut“ anstatt „seit 4 Tagen“, „bilateral“ anstatt „beide Hände“).

#### **2. (4) Beim Anamnesegespräch**

bleibe ich oft bei einer möglichen Diagnose.

fällt es mir leicht, unterschiedliche Diagnoseansätze in Betracht zu ziehen.

#### **3. (5) Während des Anamnesegesprächs,**

wenn ich dem Gedankengang des Patienten folge, verliere ich gelegentlich meinen roten Faden.

kann ich meine Ideen weiterverfolgen, selbst wenn ich dem Gedankengang des Patienten folge.

#### **4. (6) Wenn ich mich für eine Diagnose entscheiden soll,**

stört es mich nicht, meine Entscheidung auf später zu verschieben.

habe ich das Gefühl, eine Entscheidung treffen zu müssen, auch wenn ich mir nicht sicher bin.

<sup>1</sup> In Klammern steht die Nummer des Items im Originalfragebogen mit 41 Items

<sup>2</sup> Zur besseren Lesbarkeit wird im Folgenden das männliche Geschlecht verwendet. Selbstverständlich sind trotz der gewählten Form immer beide Geschlechter gemeint

**5. (7) Wenn der Patient seine Zeichen und Symptome geschildert hat,**

denke ich in seinen/ihren Worten darüber nach.

übersetze ich diese in medizinische Fachausdrücke (z.B. „Taubheitsgefühl“ wird zu „Parästhesie“).

**6. (8) In Bezug auf die allgemeine Krankengeschichte**

habe ich das Gefühl, dass ich diese oft nicht ausreichend aufnehme.

nehme ich diese zu meiner eigenen Zufriedenheit meistens ausreichend auf.

**7. (9) Während der Patient seine gesamte Krankheitsgeschichte erzählt,**

fällt es mir oft schwer, mich zu erinnern, was gesagt wurde.

schaffe ich es normalerweise, mir zu merken, was gesagt wurde.

**8. (10) Im Verlauf des Patientengesprächs stelle ich fest,**

dass mir einige Schlüsselinformationen ins Auge zu springen scheinen.

dass es oft schwierig ist, zu wissen, auf welche Informationen ich mich konzentrieren sollte.

**9. (12) In der Abwägung möglicher Diagnosen**

ziehe ich oft nicht zutreffende Diagnosen in Betracht.

bewegen sich meine Diagnosen häufig in die richtige Richtung.

**10. (13) Während ich Informationen über den Patienten sammle,**

ordnen sich diese meist von selbst in meinen Gedanken.

fällt es mir schwer, Zusammenhänge zwischen den einzelnen Informationen zu erkennen.

**11. (15) Während eines Anamnesegesprächs**

kann ich irrelevante Informationen nicht ausblenden.

fällt es mir leicht zu entscheiden, welche Information ich weglassen kann, da sie irrelevant sind.

**12. (16) Wenn die vom Patienten geschilderten Zeichen und Symptome keinen Sinn ergeben,**

dann fahre ich fort, um neue Informationen zu erhalten und eine neue Perspektive einzunehmen.

dann betrachte ich die genannten Symptome aus einer anderen Perspektive, bevor ich fortfahre weitere Informationen zu sammeln.

**13. (20) Unter Berücksichtigung der Zeichen und Symptome des Patienten**

denke ich über diese im genauen Wortlaut des Patienten nach.

denke ich an diese in möglichen Gegensätzen (z. B.: spastisch - schlaff, fortschreitend – plötzlich, unilateral - bilateral).

**14. (22) Wenn ich im Verlauf der Anamnese bereits einige Ideen über mögliche Diagnosen habe,**

geben mir neue Informationen zusätzliche Ideen.

geben mir neue Informationen keine zusätzlichen Ideen.

**15. (26) Wenn ich eine Information erhalte, die mich an eine mögliche Diagnose denken lässt,**

denke ich an frühere Informationen zurück, um abzuklären, ob diese damit zusammenpassen oder nicht.

überdenke ich selten die Informationen, die ich vorher gesammelt habe.

**16. (29) Wenn ich viel über eine bestimmte Krankheit weiß und eine Diagnose stellen muss,**

überprüfe ich die meisten Möglichkeiten, bevor ich eine Entscheidung treffe.

habe ich oft viele unterschiedliche Ideen, die ich aber nicht weiter verfolge.

**17. (30) Während mir der Fall klarer wird,**

halte ich es nicht für hilfreich, während des Prozesses die Informationen zusammenzufassen.

mache ich regelmäßig Bestandsaufnahmen der Daten und meiner Ideen.

**18. (33) Wenn ich eine allgemeine Vorstellung habe, was dem Patienten fehlen könnte,**

kann ich gewöhnlich fortfahren, bis ich eine spezifische Diagnose gefunden habe.

fällt es mir schwer, eine spezifische Diagnose zu formulieren.

**19. (34) Während des Patientengesprächs**

schaffe ich es, meine Ideen zu überprüfen, selbst wenn ich dem Patienten die Gesprächsführung überlasse.

bin ich nur dann erfolgreich, wenn ich die Richtung des Gesprächs vorgeben kann.

**20. (37) Wenn ich meine diagnostischen Ideen in Betracht ziehe, mache ich das auf der Basis**

der gesamten Informationen bis zu diesem Zeitpunkt.

einiger herausstechender Zeichen und Symptome.

**21. (38) Wenn ich nicht weiß, welche Schlüsse ich aus dem Gespräch ziehen soll,**

fällt es mir leicht, die Informationen auf eine neue Weise zu betrachten.

fällt es mir schwer, die Informationen auf eine neue Weise zu betrachten.