

Zeitschrift: Intercura : eine Publikation des Geriatriischen Dienstes, des Stadtärztlichen Dienstes und der Psychiatrisch-Psychologischen Poliklinik der Stadt Zürich

Herausgeber: Geriatriischer Dienst, Stadtärztlicher Dienst und Psychiatrisch-Psychologische Poliklinik der Stadt Zürich

Band: - (2000-2001)

Heft: 70-71

Artikel: Reliabilität und Validität des TFDD : Test zur Früherkennung von Demenz mit Depressionsabgrenzung

Autor: Ihl, R. / Gross-Kapanke, B.

DOI: <https://doi.org/10.5169/seals-790051>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 15.02.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Reliabilität und Validität des TFDD

Test zur Früherkennung von Demenz mit Depressionsabgrenzung

von R. Ihl und B. Gross-Kapanke

Prozedere

Die Untersuchungen zur Demenzfrüherkennung und zur Depressionsabgrenzung wurden sequentiell durchgeführt. Für beide Testteile wurde jeweils eine separate Gruppe getrennt untersucht und ausgewertet. Soweit sich Patienten zu einer wiederholten Testung bereit erklärten, wurden diese in die Test-Retest-Validitätsberechnung eingeschlossen.

Versuchspersonen

Für die Untersuchung des Demenz Screenings wurden 88 Patienten mit Alzheimerkrankheit, 52 Patienten mit Depression und 37 gesunde altersvergleichbare Kontrollpersonen untersucht. Um eine möglichst homogene Patientengruppe von praktischer Relevanz zu erhalten, wurden ausschliesslich Patienten mit wahrscheinlicher Alzheimerkrankheit (NINCDS/ARDA-Kriterien, (McKhann, Folstein et al., 1984), nachfolgend als "Patienten mit Alzheimerkrankheit" bezeichnet) eingeschlossen, da diese den grössten Anteil unter den Demenzkranken ausmachen. 88 Patienten mit Alzheimerkrankheit und 52 mit depressiver Störung (ICD-10, F31.3-31.5, F32 und F33) wurden untersucht (Tab. 1). 37 altersvergleichbare gesunde Kontrollpersonen wurden zusätzlich eingeschlossen (Alter >60 Jahre). Einschlusskriterium waren Kooperationsbereitschaft und die Einwilligung nach Aufklärung über Sinn und Zweck der Untersuchung. Ausschlusskriterien waren anders verursachte Störungen der Kognition (z.B. Exsikkose, Hirntumor) sowie das Vorliegen anderer psychiatrischer Erkrankungen.

Für die Untersuchung des Depressionsteils wurden 18 Patienten mit Alzheimerkrankheit, 18 Patienten mit depressiver Störung und 10 Kontrollpersonen nach den gleichen Einschlusskriterien untersucht (Tab. 1).

Tabelle 1: Versuchspersonen und Patientencharakteristika

Parameter	Alzheimer- krankheit	depressive Störung	Kontrollper- sonen
n (Demenz Screening) = 177	88	52	37
Alter (Median)	81	71	56
Geschlecht (m/w)	15/73	15/37	18/19
Schulbildung	8	8	10
n (Depressionsteil) = 46	10	18	18
Alter (Median)	79.5	70	77.5

Um die Konvergenzvalidität zu ermitteln, wurde die Geriatric Depression Scale (GDS, Yesavage et al 1983) ebenfalls durchgeführt.

Demenzitems

Sensitivität/Spezifizität

Die Daten der Untersuchungsgruppen wurden mit dem Mann-Whitney-U-Test auf Gruppenunterschiede geprüft. Es wurden alle drei möglichen Gruppenvergleiche berechnet. Das bei gerichteter Hypothese einseitige Signifikanzniveau von $p < 0.05$ wurde mittels Bonferronikorrektur bei drei Tests auf $p < 0.016$ adaptiert.

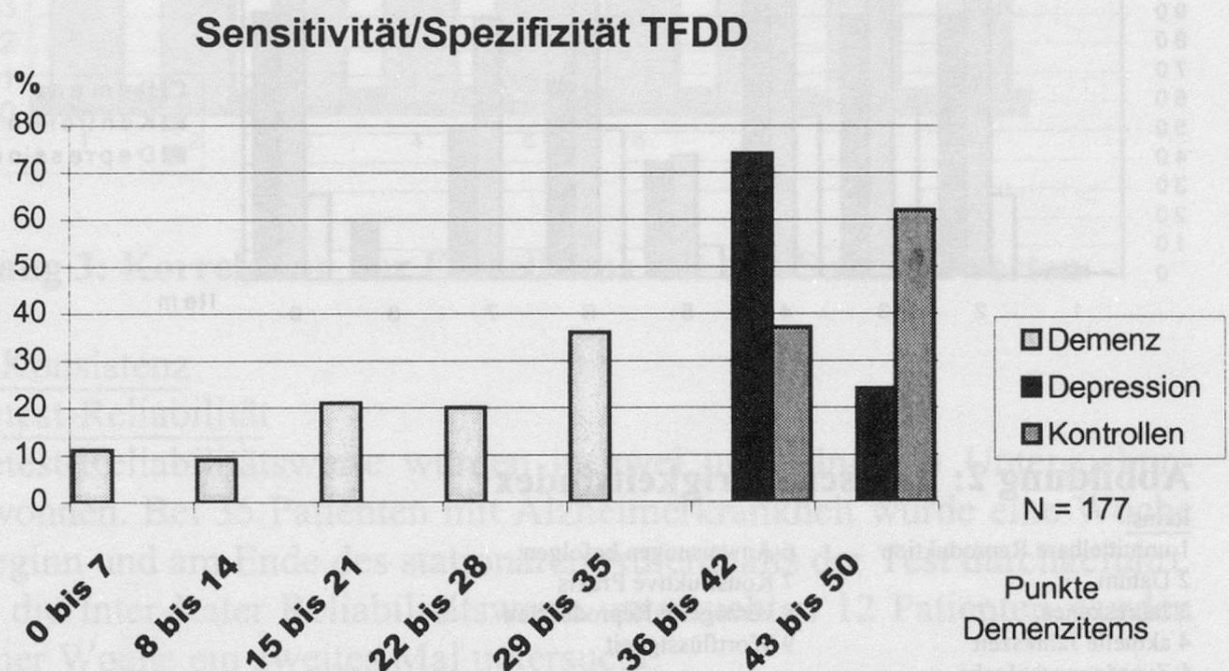


Abbildung 1: Sensitivität/Spezifizität Demenzitems

Die Gruppenunterschiede erwiesen sich im Mann-Whitney-U-Test als signifikant für den Vergleich Kontrollpersonen vs. Patienten mit Alzheimerkrankheit ($n = 125$, $U = 0$, $p < 0,001$), ebenso für Patienten mit Depression vs. Patienten mit Alzheimerkrankheit ($n = 140$, $U = 0$, $p < 0,001$) und für Kontrollpersonen vs. Patienten mit Depression ($n = 89$, $U = 485,5$, $p < 0,001$). Der höchste erreichte Punktwert bei Patienten mit Alzheimerkrankheit lag bei 35 Punkten. Die Kontrollgruppe erreichte Werte zwischen 38

und 49 Punkten. Es fand sich damit keine Überlappung beider Gruppen (Sensitivität und Spezifität 100%, Abb. 1).

Depressive Patienten erreichten Punktwerte zwischen 36 und 47 Punkten. Hier fand sich ebenso keine Überschneidung mit den Werten der Patienten mit Alzheimerkrankheit (Spezifität 100 %, Abb. 1). Kontrollpersonen erreichten trotz hohen Überlappungsbereichs signifikant höhere Werte als Patienten mit Depression.

Anteil vollständig richtiger Lösungen pro Item

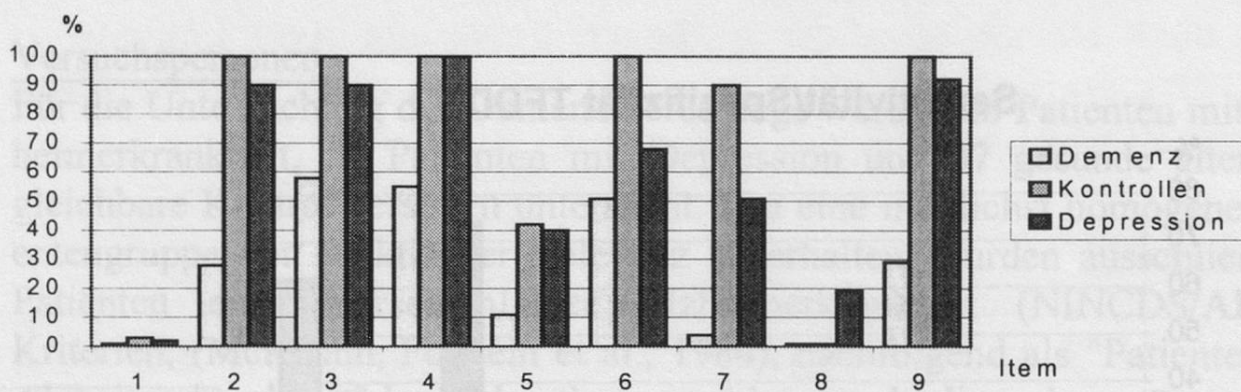


Abbildung 2: Itemschwierigkeitsindex

Items:

- | | |
|-----------------------------|---------------------------|
| 1 unmittelbare Reproduktion | 6 Anweisungen befolgen |
| 2 Datum | 7 Konstruktive Praxis |
| 3 Jahreszeiten | 8 verzögerte Reproduktion |
| 4 aktuelle Jahreszeit | 9 Wortflüssigkeit |
| 5 Zuordnungsaufgabe | |

Der Schwierigkeitsgrad der einzelnen Items ist in Abb. 2 dargestellt. Es gelang mit allen Items, bei den vollständig richtigen Lösungen einen Unterschied zwischen Patienten mit Alzheimerkrankheit und gesunden Kontrollpersonen aufzuzeigen (durchgängig geringerer Lösungsprozentsatz bei Patienten mit Alzheimerkrankheit). Die Unterschiede fielen am grössten bei Item 2 "Datum", Item 7 "Uhrtest" und Item 9 "Wortflüssigkeit" aus.

Zur Untersuchung der Item-Reliabilität wurde die Korrelation der Einzeli-tems mit dem Testsummenwert und Cronbachs α berechnet. Der Beitrag, den jedes Item zum Test liefert, wurde durch Unterdrücken des jeweiligen Items bestimmt. Alle Items korrelieren mit dem Gesamtwert $r = 0.5$ und höher (siehe Abbildung 3).

Cronbachs α war durchgängig höher als 0.8 für die einzelnen Items und betrug 0.8803 für die gesamten Demenzitems.

Korrelation der Einzelitems mit den Summenwerten

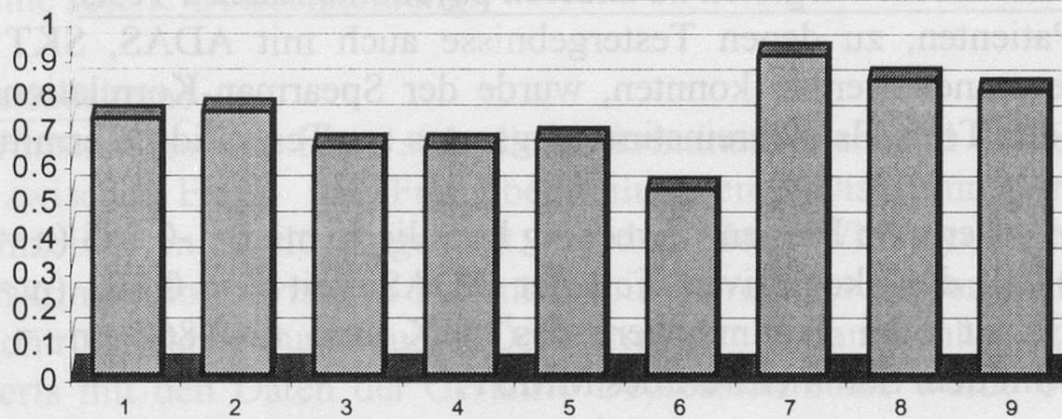


Abbildung 3: Korrelation der Einzelitems mit den Summenwerten

Interne Konsistenz

Test-Retest-Reliabilität

Test-Retest-Reliabilitätswerte wurden in zwei unabhängigen Untersuchungen gewonnen. Bei 35 Patienten mit Alzheimerkrankheit wurde eine Woche nach Beginn und am Ende des stationären Aufenthalts der Test durchgeführt. Die für die Inter-Rater Reliabilitätswerte untersuchten 12 Patienten wurden nach einer Woche ein zweites Mal untersucht.

Die Daten der Patienten dieser Untersuchung waren auch die Basis für die Berechnung eines Schwierigkeitsindex der Items sowie deren Item-Reliabilität.

Bei der Untersuchung zur Test-Retest-Reliabilität fand sich ein r_s von 0.868 ($p < 0.001$, $n = 35$) in der Gruppe mit variablem Untersuchungsintervall (Spannweite 2 – 12 Wochen). Bei der Untersuchung der Gruppe mit konstantem Abstand von einer Woche fand sich eine höhere Test-Retest-Reliabilität von $r_s = .99$ ($p < 0.001$, $n = 12$) für Patienten mit Demenz.

Inter-Rater-Reliabilität

Zur Prüfung der Inter-Rater-Reliabilität wurden 12 Patienten durch zwei trainierte Raterinnen untersucht. Dabei wurde wie folgt vorgegangen: je 9

Patienten mit Depression und Demenz wurden eingeschlossen (Selektionskriterien siehe oben). Jeweils die Hälfte der Patienten wurden durch Raterin 1, bzw. Raterin 2 befragt. Beide Raterinnen führten die Bewertung selbständig und parallel durch.

Die Interrater Reliabilität betrug r_s 0.993 ($p < 0.001$, $n = 12$).

Externe Validität im Vergleich zu anderen psychometrischen Tests.

Für die Patienten, zu denen Testergebnisse auch mit ADAS, SKT oder MMST gewonnen werden konnten, wurde der Spearman-Korrelationswert zwischen den Tests als Übereinstimmungsmass zur Testvalidität ermittelt ($n = 177$).

Die Summenwerte im Demenz Screening korrelierte mit $r_s = -0.888$ ($n = 107$, $p < 0.001$) mit dem kognitiven Teil der ADAS, mit $r_s = 0.862$ ($n = 110$, $p < 0.001$) mit dem Summenwert des STK und $r_s = 0.862$ ($n = 110$, $p < 0.001$) mit dem Summenwert des MMST.

Spezifizität im Vergleich zum Mini-Mental-Status-Test

Für die Teilgruppe der Untersuchungspersonen, die auch den Mini-Mental-Status-Test (MMST) absolvierten, sind in Abbildung 4 die gefundenen Punktwerte nach Diagnosen dargestellt.

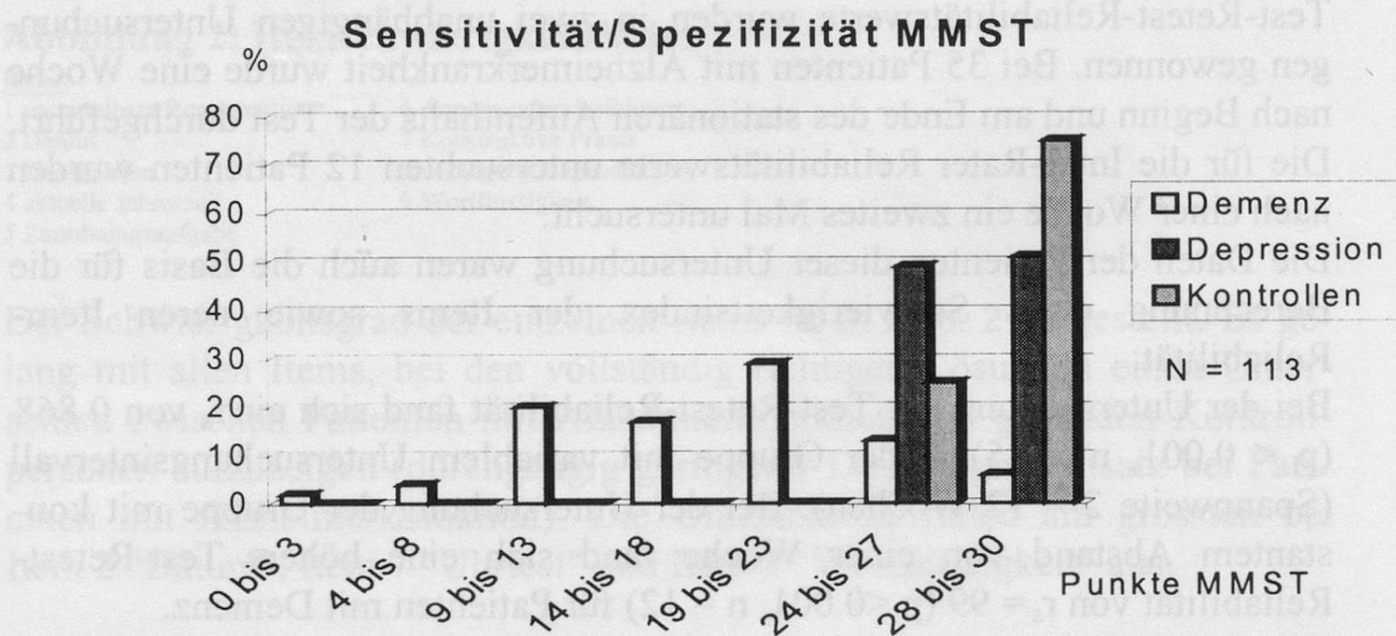


Abbildung 4: Sensitivität und Spezifizität des MMST

Im Vergleich mit Abbildung 1 wird deutlich, dass beide Tests Patienten mit Depression Werte oberhalb des cut-off Werts für Demenz zuordnen. Der MMST ordnet allerdings 13 Patienten mit Alzheimerkrankheit in den Bereich gesunder Kontrollpatienten und von Patienten mit Depression ein. Leicht Demenzkranke werden so nicht identifiziert, die Abgrenzung depressiver Pseudodementen erschwert. Die hier erhobenen Daten zum TFDD zeigen keine solche Überschneidung.

Depressionsitems

In Voruntersuchungen hatte sich die Diskrimination anhand des Summenwerts zwischen Eigen- und Fremdbeurteilung einer Diskrimination anhand von Einzelwerten überlegen gezeigt. Mit dem Mann-Whitney-U-Test wurde die Diskrimination zwischen Patienten mit depressiver Störung und gesunden Kontrollpersonen untersucht. Der korrelative Zusammenhang des Summenwerts mit den Daten der Geriatric Depression Scale wurde berechnet, um einen Hinweis auf die Konvergenzvalidität zu gewinnen.

Die *Korrelation* des Summenwertes auf Fremd- und Selbstbeurteilung mit dem Summenwert der *Geriatric Depression Scale* betrug $r_s = 0.73$ ($n = 33$, $p < 0.01$). Patienten mit Depression konnten signifikant von gesunden Kontrollpersonen abgegrenzt werden ($n = 36$, $u = 0$, $p < 0.001$). (Der höchste Wert für Kontrollpersonen lag bei 7 Punkten, der niedrigste für Patienten mit depressiver Störung bei 9 Punkten. Selbst- und Fremdbeurteilung der Depressivität korrelierten mit $r_s = 0.869$ ($n = 46$, $p < 0.01$).

Die Inter-Rater-Reliabilität lag in der Test-Retest-Untersuchung zu Beginn bei $r_s = 0.753$. ($n = 18$, $p < 0.001$) und in der Folgeuntersuchung bei $r_s = 0.997$. ($n = 10$, $p < 0.001$). Die Test-Retest-Reliabilität erreichte bei deutlich reduzierten n $r_s = 0.7$ für Raterin 1 und $r_s = 0.651$ für Raterin 2 ($n = 8$, $p < 0.05$ für beide Werte).