

Beobachten statt testen

Einschätzung in der Unterstützten Kommunikation: warum Normtests hier scheitern — und was stattdessen trägt

Lars Tiedemann

Dipl. Heilpädagoge (FH) · AssistUK

5. August 2026

Arbeitsfassung. »Nicht testbar« steht in vielen Akten — als wäre das eine Eigenschaft des Menschen und nicht des Tests. Dieser Text begründet, warum normierte Verfahren bei Menschen mit komplexen Kommunikationsbedürfnissen systematisch versagen, und beschreibt die Alternative, die unser Feld entwickelt hat: dynamisch einschätzen, kriteriumsorientiert einordnen, individuell Verlauf messen. Er ist die fachliche Grundlage hinter unserer Einschätzungspraxis und dem Wortzugriffs-Werkzeug. Quellen mit Prüfvermerk am Ende.

Inhalt

1. »Nicht testbar«
 2. Warum Normtests hier systematisch scheitern
 3. Der Wechsel der Frage
 4. Dynamisch einschätzen: Testen, Lehren, wieder Testen
 5. Einordnen ohne Normen: die Landkarten
 6. Verlauf statt Vergleich: die Person als ihr eigener Maßstab
 7. Beobachtung, die den Namen verdient
 8. Wie wir es bei AssistUK halten
 9. Grenzen dieses Textes
 10. Fazit in fünf Sätzen
 11. Literatur (mit Prüfvermerk)
-

1. »Nicht testbar«

Ich habe das Wort in unzähligen Berichten gelesen, und es hat mich nie losgelassen: *nicht testbar*. Gemeint ist: Das Kind hat auf die Testkarten nicht gezeigt, die Aufgaben nicht bearbeitet, die Mitarbeit »verweigert«. Schlussfolgert wird daraus — mal vorsichtig, mal nicht — ein sehr niedriges Fähigkeitsniveau. Und auf dieser Grundlage werden dann Schullaufbahnen, Vokabulargrößen und Zutrauen bemessen.

Dabei beschreibt »nicht testbar« gar nicht das Kind. Es beschreibt die Begegnung zwischen einem Menschen, dessen Ausdruckswege blockiert sind, und einem Verfahren, das genau diese Ausdruckswege voraussetzt. Der Test misst bei dieser Begegnung nicht Kompetenz — er misst Zugänglichkeit. Und die ist keine Eigenschaft der Person, sondern der Situation.

2. Warum Normtests hier systematisch scheitern

Normierte Tests beruhen auf drei Annahmen, und bei Menschen mit komplexen Kommunikationsbedürfnissen bricht jede einzelne.

Erstens: Die Antwort muss den Kopf verlassen können. Standardisierte Verfahren verlangen Zeigen, Sprechen, Hantieren — in Normzeit. Wer motorisch nicht zeigen, nicht sprechen, nur verzögert reagieren kann, »versagt« unabhängig davon, was er weiß. Der Cerebralparese- und der Rett-Text buchstabieren das im Detail: Die Latenz ist Teil der Behinderung, der Test wertet sie als Nichtwissen.

Zweitens: Die Normstichprobe muss die Person enthalten. Normwerte sagen, wo jemand im Vergleich zu Gleichaltrigen steht, die so ähnlich sind wie die Eichstichprobe. Kinder, die mit Augensteuerung antworten, mit Symbolen sprechen, mit Assistenz leben, waren in keiner Eichstichprobe. Der Vergleich vergleicht mit einer Gruppe, zu der die Person nie gehörte.

Dritens: Die Testsituation muss die Fähigkeit hervorlocken. Fremder Raum, fremde Person, sinnfreie Aufgaben, Zeitdruck — für viele unserer Zielgruppen ist das die maximal ungünstige Bedingung: Was im vertrauten Spiel sichtbar wäre, verschwindet am Testtisch. Gemessen wird die Leistung unter den schlechtesten Umständen und ausgegeben als »das Niveau«.

Die Folge ist keine zufällige Streuung, sondern ein **systematischer Fehler in eine Richtung: nach unten**. Jeder der drei Brüche drückt das Ergebnis — keiner hebt es. Wer das weiß, liest Testberichte über nicht sprechende Menschen anders: als Untergrenze mit unbekanntem Abstand zur Wahrheit.

3. Der Wechsel der Frage

Die Alternative beginnt nicht bei anderen Instrumenten, sondern bei einer anderen Frage. Der Normtest fragt: *Wo steht diese Person im Vergleich zu anderen?* Die Einschätzung in der Unterstützten Kommunikation (UK) fragt: **Was gelingt dieser Person unter welchen Bedingungen — und was wäre der nächste Schritt?**

Das ist kein Trostpreis, sondern die fachlich ergiebige Frage. Denn für die Praxis ist der Prozentrang herzlich egal; handlungsleitend ist, *unter welchen Bedingungen* mehr gelingt: mit längerer Wartezeit? Mit größeren Feldern? Am Vormittag? Mit der vertrauten Assistentin? Nach einem Modell? Genau diese Bedingungs-Informationen wirft der Normtest weg (er standardisiert sie ja gerade fort) — und genau sie sind das Gold der Förderplanung.

4. Dynamisch einschätzen: Testen, Lehren, wieder Testen

Das wichtigste Verfahren dieser anderen Logik heißt **dynamisches Assessment**, und Martha Snell hat es für unser Feld schon 2002 programmatisch begründet: Statt einmal statisch zu messen, was jemand *heute allein* kann, prüft man in einer Schleife — anbieten, kurz lehren (ein Modell, eine Hilfe, eine veränderte Bedingung), wieder anbieten. Gemessen wird die **Veränderung**: Wie viel Unterstützung braucht es, damit etwas gelingt? Was nimmt die Person aus einem Modell mit?

Der Gewinn ist doppelt. Erstens misst man damit die *Lernfähigkeit* statt des Ist-Bestands — also genau das, was für Prognose und Systemwahl zählt (die Zone der nächsten Entwicklung, um den klassischen Begriff zu nennen). Zweitens ist das Verfahren fair gegenüber allem, was der Normtest bestraft: Zeit wird gegeben statt gestoppt, Wege werden angepasst statt vorgeschrieben, Vertrautes ist erlaubt statt verboten. Unsere Wortzugriffs-Einschätzung ist exakt so gebaut — Übung vorweg, alle drei Wege fair angeboten, Abbruch ist ein Ergebnis und kein Versagen —, und die Einschätzungen in unseren Zugangs-Plattformen folgen derselben Linie: dieselben Aufgaben, dokumentierte Bedingungen, Wiederholung nach Monaten.

5. Einordnen ohne Normen: die Landkarten

»Keine Normen« heißt nicht »keine Ordnung«. Das Feld hat kriteriumsorientierte Landkarten entwickelt, die beschreiben, *wo auf einem Entwicklungsweg* jemand steht — ohne zu behaupten, wo er im Vergleich zur Alterskohorte stehen müsste. Die für uns wichtigsten: die **Communication Matrix** (Rowland) für die frühen Ebenen — vom vorintentionalen Verhalten bis zur Symbolkombination; das **DAGG-3** (Tobii Dynavox mit Vicki Clarke) für die Stufen kommunikativer Kompetenz entlang der Light-Domänen; und für die ganz frühen, körpernahen Signale die Systematik von Coupe O'Kane und Goldbart. Solche Instrumente beantworten die Frage »Was ist der nächste Schritt?« direkt — deshalb koppeln wir an sie unsere Vokabular-Empfehlungslogik (der Zugriff-Text beschreibt das).

Zwei Redlichkeits-Hinweise gehören dazu. Diese Landkarten sind Ordnungsrahmen, keine geeichten Messinstrumente — sie machen Beobachtung besprechbar, sie ersetzen sie nicht. Und wie Fachlichkeit die Einschätzung prägt, ist selbst untersucht: Erfahrene UK-Spezialistinnen gehen nachweislich anders vor als Generalistinnen (Dietz et al. 2012) — Einschätzung ist Expertise, kein Formular.

6. Verlauf statt Vergleich: die Person als ihr eigener Maßstab

Wo der Gruppenvergleich wegfällt, übernimmt der **Selbstvergleich über Zeit** — und er ist strenger, als er klingt, wenn man ihn ordentlich macht: dieselben Situationen, dieselben Aufgaben, dokumentierte Bedingungen, feste Abstände. Dann werden Sätze möglich, die kein Normtest liefern kann: *Vor sechs Monaten zwei Felder mit 2000 Millisekunden — heute acht Felder mit 900. Damals nur auf Ansprache — heute dreimal pro Stunde selbstinitiiert.*

Für individuelle Ziele hat sich ergänzend das **Goal Attainment Scaling** bewährt — vorab wird je Ziel eine fünfstufige Skala vom »weniger als erwartet« bis »viel mehr als erwartet« definiert, und der Fortschritt wird daran abgelesen. Auch hier gilt die kritische Fußnote (Schlosser 2004): GAS misst Zielerreichung, nicht Ursache — es ersetzt keine kontrollierte Wirksamkeitsprüfung. Aber als gemeinsame, transparente Verabredung, woran alle den Fortschritt erkennen wollen, ist es dem vagen »macht Fortschritte« des Alltags weit überlegen. Und gegenüber Kostenträgern ist der dokumentierte Selbstvergleich die Währung, die zählt: nachlesbare Entwicklung statt behaupteter.

7. Beobachtung, die den Namen verdient

Bleibt das Wort, das im Titel steht — denn »wir beobachten erst mal« kann auch eine Ausrede fürs Nichtstun sein. Beobachtung, die den Namen verdient, hat drei Merkmale:

Sie ist situiert. Beobachtet wird da, wo die Person lebt und mag — im Spiel, beim Essen, im Streit. Nicht, weil es netter ist, sondern weil Kompetenz kontextgebunden ist: Die interessanteste Frage ist fast immer, *wo* etwas gelingt und *wo* nicht.

Sie ist verschriftlicht. Datum, Situation, Signal, Reaktion, Bedingung — kurz, aber schriftlich. Das Gedächtnis glättet; die Notiz bewahrt die Ausreißer, und die Ausreißer nach oben sind die wichtigsten Daten der ganzen Einschätzung (sie zeigen, was unter günstigen Bedingungen geht).

Sie mündet in eine Hypothese. Gute Beobachtung endet nicht mit »wir beobachten weiter«, sondern mit einem prüfbaren Satz: *Wenn wir die Wartezeit verdoppeln, steigt die Trefferquote.* Dann wird die Bedingung geändert, wieder beobachtet — und aus Beobachtung wird dynamisches Assessment. Der Kreis schließt sich.

8. Wie wir es bei AssistUK halten

Unsere Regeln, offen gelegt: Jede unserer Einschätzungen — die Wortzugriffs-Einschätzung, die Beobachtungsstrecken in den Zugangs-Plattformen — trägt den ausdrücklichen Vermerk, **kein normiertes Testverfahren** zu sein: kein Altersvergleich, keine Prozentränge, Vergleich nur mit sich selbst über die Zeit. Alle Aufgaben sind an die Zugangswege anpassbar, Wartezeiten einstellbar, Abbrüche sind dokumentierte Ergebnisse. Berichte nennen die Bedingungen mit (Tageszeit, Zugangsweg, Begleitung), weil eine Zahl ohne Bedingungen keine Information ist. Und das Wort »Diagnostik« benutzen wir für unsere eigenen Angebote nicht — nicht aus Bescheidenheit, sondern aus Präzision: Diagnosen stellen andere; wir schätzen ein und beobachten, und wir sagen dazu, was das leistet und was nicht.

9. Grenzen dieses Textes

Dieser Text kritisiert Normtests für eine bestimmte Zielgruppe — nicht Normtests überhaupt: Wo Zugangswege intakt sind und die Person der Eichstichprobe entspricht, haben standardisierte Verfahren ihren Wert, auch begleitend in der UK (etwa fürs Sprachverständnis mit angepasstem Antwortmodus). Die Kunst ist die Anpassung, und über angepasste Durchführung muss ein Bericht ehrlich Auskunft geben.

Dynamisches Assessment und kriteriumsorientierte Landkarten sind fachlicher Konsens und plausibel begründet; ihre Messgüte im strengen Sinn (Reliabilität, Prognosekraft) ist deutlich dünner erforscht als die der klassischen Tests — man tauscht bekannte Schwächen gegen weniger vermessene. Ich halte den Tausch für richtig und sage trotzdem dazu, dass es ein Tausch ist.

10. Fazit in fünf Sätzen

»Nicht testbar« beschreibt nie den Menschen, sondern die Begegnung zwischen blockierten Ausdruckswegen und einem Verfahren, das genau sie voraussetzt — der Fehler zeigt systematisch nach unten. Die bessere Frage ist nicht »Wo steht sie im Vergleich?«, sondern »Was gelingt unter welchen Bedingungen — und was ist der nächste Schritt?«. Dafür hat

unser Feld das Handwerk: dynamisch einschätzen (anbieten, lehren, wieder anbieten), kriteriumsorientiert einordnen (Communication Matrix, DAGG-3), und den Verlauf der Person gegen sich selbst messen — gleiche Aufgaben, dokumentierte Bedingungen, feste Abstände. Beobachtung verdient ihren Namen, wenn sie situiert, verschriftlicht und hypothesengeleitet ist. Und über allem steht die Konsequenz aus dem systematischen Fehler: Im Zweifel spricht die Einschätzung für die Person, nicht gegen sie.

11. Literatur (mit Prüfvermerk)

Alle Quellen am 05.08.2026 auf Existenz und Kernaussage geprüft (✓); GAS-Kritik und DAGG-Angaben aus früheren geprüften Beständen übernommen.

1. ✓ **Snell, M. E. (2002).** Using dynamic assessment with learners who communicate nonsymbolically. *Augmentative and Alternative Communication*, 18(3), 163–176. — Programmatik des dynamischen Assessments: Testen–Intervenieren–Retesten statt statischer Messung.
2. ✓ **Dietz, A., Quach, W., Lund, S. K. & McKelvey, M. (2012).** AAC assessment and clinical-decision making: The impact of experience. *Augmentative and Alternative Communication*, 28(3), 148–159. — Einschätzungsvorgehen unterscheidet sich deutlich nach Spezialisierung; Einschätzung als Expertise.
3. ✓ **Rowland, C. (1996/2004).** *Communication Matrix*. Oregon Health & Science University. — Kriteriumsorientierte Landkarte der frühen Kommunikationsentwicklung.
4. ✓ **DAGG-3 — Dynamic AAC Goals Grid, 3. Ausgabe (2023).** Tobii Dynavox mit Vicki Clarke. — Kriteriumsorientierte Stufen entlang der Light-Domänen; Ist-Stand → Ziele → Verlauf; keine Normtabellen. (Geprüft 12.07.2026.)
5. ✓ **Coupe O'Kane, J. & Goldbart, J. (1998).** *Communication Before Speech: Development and Assessment* (2. Aufl.). London: David Fulton. — Systematik und Einschätzung vorsymbolischer Kommunikation.
6. ✓ **Schlosser, R. W. (2004).** Goal attainment scaling as a clinical measurement technique in communication disorders: A critical review. *Journal of Communication Disorders*, 37(3), 217–239. — GAS: nützlich als Zielverabredung, kein Wirksamkeitsnachweis. (Geprüft 16.07.2026.)
7. ✓ **Light, J. & McNaughton, D. (2014).** Communicative competence for individuals who require AAC. *Augmentative and Alternative Communication*, 30(1), 1–18. — Die Domänen, entlang derer eingeordnet wird.
8. ✓ **Tiedemann, L. (2026).** Vom Zugriff zum Vokabular (Kopplung Einschätzung→Empfehlung) · »Wir sprechen hier alle ...«? (Kompetenzvermutung, Zone der nächsten Entwicklung) · Der Umweg ist der Weg · Die Augen antworten (Latenz und Bedingungen). AssistUK-Grundlagenreihe. — Nachbartexte, deren Einschätzungslogik dieser Text begründet.

Die fachliche Position, Auswahl und Verantwortung liegen bei mir. — L. T.