

Konzentrationstest GU

[Aus 100 zweistelligen Zahlen möglichst schnell diejenigen 25 Zahlen identifizieren, deren erste Ziffer gerade und deren zweite Ziffer ungerade ist]

Der GU erfüllt alle notwendigen Voraussetzungen eines Konzentrationstests im Sinne von Westhoff & Hagemeyer (2005). Hierzu gehören geringe Anforderungen an die Wahrnehmung, das Gedächtnis, die Lernfähigkeit, die Problemlösefähigkeit und die Intelligenz. Hohe Konzentration zeigt sich darin, die Aufgaben möglichst zügig, hinreichend ausdauernd und zugleich möglichst Fehler frei zu bearbeiten. In diesem Sinne kann der Test als inhaltsvalide angesehen werden.

Ein Testdurchgang des GU umfasst jeweils 100 stets nach bestimmten Zufallsprinzipien ausgewählte zweistellige Zahlen. 25 dieser Zahlen (Zielzahlen) beginnen mit einer geraden und enden mit einer ungeraden Ziffer (z.B.: 85 in Abbildung 1). Die verbliebenen 75 Zahlen dienen als Distraktoren. Die Zielzahlen befinden sich stets an Zufallspositionen und sollten vom Probanden so schnell wie möglich, aber dennoch korrekt mit linker Maustaste angeklickt werden. (weitere Angaben bei Jacobs 2013)

Abbildung 1: Beispiel für einen möglichen Testdurchgang im GU

**Erste Ziffer gerade
zweite Ziffer ungerade!**

52	92	85	23	95	82	92	56	58	47
31	31	85	46	64	97	72	94	32	82
59	46	47	74	65	62	86	54	61	69
76	82	26	57	49	37	93	66	91	18
18	97	65	96	19	41	46	84	29	52
22	51	88	67	48	96	44	54	42	75
68	57	63	67	72	88	64	13	94	92
74	63	23	84	81	31	48	14	58	97
38	96	79	72	96	68	89	83	24	72
82	84	19	85	87	85	66	45	53	23

Bearbeitung bestätigen!

Der Konzentrationstest GU basiert auf einer alten Computerfassung von Jacobs (2002a,b), wurde später leicht modifiziert mit einer verbesserten Anweisungsinstruktion versehen und auf insgesamt 3 Testdurchgänge verlängert (Jacobs 2013). Mit dieser zweiten Version fanden dann einige empirische Analysen statt (Jacobs 2013,2014). Hierbei fiel jedoch auf, dass manche Probanden unplausible und zum Teil unbrauchbare Daten produzierten, die auf weiterhin bestehende Verständnisprobleme hindeuteten. Verbesserungen der nunmehr dritten Version sehen wie folgt aus:

1. Vor der eigentlichen Testung wird ein auf 16 Zahlen beschränkter Probedurchgang eingeführt. Er muss notfalls mit veränderten Zahlen sooft wiederholt werden, bis er absolut fehlerfrei absolviert werden kann.
2. Ca. 1.5 Sekunden vor jedem Testdurchgang erhält die Testperson einen Hinweis, sie solle sich darauf vorbereiten, dass nun ein Testdurchgang (z.B. Testdurchgang 2) bevorstehe.

Diese Maßnahmen führten dann zu keinem Ausfall mehr, der mit mangelndem Instruktionsverständnis hätte in Verbindung gebracht werden können. Lediglich ein Proband von 99 TeilnehmerInnen produzierte ungültige Daten.

Als zentrale Messvariable der Konzentrationsfähigkeit gilt die durchschnittliche Bearbeitungszeit (Arbeitstempo) eines Testdurchgangs in Sekunden. In den ersten Versionen wurde zusätzlich noch der Prozentsatz der erkannten Zielitems (analog dem Auslassungsfehler) ermittelt, dem wegen mäßiger Reliabilität aber keine weitere Bedeutung beigemessen wurde.

Tabelle 1 stellt die wichtigsten bisher vorliegenden deskriptiven Ergebnisse des GU zusammen. Die ersten beiden Studien basieren auf der zweiten Version des GU. Im Wintersemester 2015/16 hat der Verfasser mit den Studierenden seiner Seminare eine weitere Studie mit der nun verbesserten dritten Version des GU durchgeführt, deren Ergebnisse ebenfalls in Tabelle 1 enthalten sind.

Tabelle 1 Deskriptive Ergebnisse zum GU: Durchschnittliche Bearbeitungszeit aus 3 Testdurchgängen in Sekunden

	α	M	s	Md	Schiefe	Kurtosis	N
Jacobs 2013	.92	96.9	22.3	93.3	.80	.41	102
Jacobs 2014	.90	101.9	26.2	97.2	.87	.25	62
WS 2015/16	.97	95.7	23.4	90.5	.80	-.01	98

Wie aus Tabelle 1 hervorgeht, erzielten die einzelnen Stichproben annähernd vergleichbare Mittelwerte und Standardabweichungen und die Testwerte folgen ziemlich einheitlich bis auf eine leicht linkssteile Drift weitgehend einer Normalverteilung. Die Reliabilitäten (Cronbachs α) fallen recht ordentlich und in der letzten Version sogar hervorragend aus. Systematische Retestanalysen stehen aber noch aus. Lediglich in der Studie von Jacobs (2014) wiederholte ein Teil der Probanden den GU im Abstand von 8 bis 12 Tagen. Der hierbei ermittelte Retestkoeffizient von $r_{tt} = .83$ fällt ebenfalls zufriedenstellend aus, basiert allerdings nur auf einem N von 32 Probanden. Als Paralleltest bietet sich der UG an, bei dem die erste Ziffer ungerade und die zweite Ziffer gerade ist. Als weiterer Paralleltest könnte der GG_UU dienen. Er verlangt die Identifikation von Zahlen, die nur gerade oder nur ungerade Ziffern enthalten. Es werden entsprechende Versionen angeboten, aber es fehlen bisher empirische Analysen dazu.

Überlegungen auf der Basis der neuen Version des GU (2015)

Berücksichtigung von Fehlern in einem Konzentrationsleistungsmaß

Bearbeitungszeiten sind natürlich wenig aussagekräftig, wenn sie quasi auf Kosten der Genauigkeit zustande kommen. Krasse Ausreißer im Hinblick auf Bearbeitungszeiten oder Fehler wurden in den empirischen Analysen vorsorglich ausgeschlossen. Bei den verbliebenen Studierenden war dann auch keine nennenswerte Korrelation zwischen Bearbeitungszeit und Auslassungsfehler mehr festzustellen. Allerdings blieb der Verwechslungsfehler unberücksichtigt, unter anderem, weil er relativ selten auftrat. Um aber dennoch ein fundiertes Maß für eine fehleradjustierte Bearbeitungszeit zu erstellen, wurde als weiteres Konzentrationsmaß die Konzentrationsleistung (KL) als Anzahl der Fehler adjustierten korrekten Angaben pro Minute wie folgt erfasst:

$KL = (\text{Anzahl richtig} - \text{Anzahl ausgelassen} - \text{Anzahl verwechselt}) * 60 / \text{Bearbeitungszeit in Sekunden}$.

Auch dieses Maß bezieht sich auf den Durchschnitt der Testdurchgänge. Bereits bei Jacobs (2014) wurde ein Leistungsmaß (Anzahl richtig/Bearbeitungszeit) erfasst. Allerdings sind dabei Verwechslungsfehler gar nicht und Auslassungsfehler nur teilweise berücksichtigt. Mit Hilfe der neuen KL werden fragwürdige Daten (z.B. hohe Geschwindigkeit bei vielen Fehlern) nun vollständig korrigiert und Verfälschungen direkt transparent (z.B. negativer Konzentrationsleistungswert). Konzentrationsleistung und Bearbeitungszeit korrelieren immer noch $r = -.85$ miteinander. Sie messen also weitgehend dasselbe, weil die Konzentrationsleistung im Wesentlichen von der Bearbeitungsgeschwindigkeit bestimmt wird und Fehler infolge ihrer relativ geringen Anzahl ein niedrigeres Gewicht haben.

Reliabilität und Testlänge

Aus den Mittelwerten in Tabelle 1 lässt sich folgern, der durchschnittliche Studierende benötige für den GU eine reine Bearbeitungszeit von fast 5 Minuten ($M * 3$ Durchgänge). Rechnet man Instruktion und Probedurchgang hinzu, so erfordert die Gesamttestung bei vielen Studierenden einen ziemlich hohen Zeitaufwand. Die Testung selbst wird von den Studierenden auch als recht anstrengend und lästig empfunden (Jacobs 2015b, S. 20), was zum Teil auch der Dauer des Tests geschuldet ist.

Aus diesem Grunde werden hier Überlegungen angestellt, welche Konsequenzen eine Verkürzung des Tests nach sich ziehen würde. Hierbei wird zunächst die Beziehung zwischen Testlänge und Reliabilität ermittelt und geschätzt. Um Cronbachs α für 2 Durchgänge empirisch zu ermitteln, wurden Reliabilitätsanalysen mit den ersten beiden Testdurchgängen durchgeführt. Die nach Spearman-Brown erwartete Reliabilität für 2 Durchgänge ergibt sich dadurch, dass die Aufgabenanzahl von 3 auf 2 Durchgänge, also um ein Drittel gekürzt wurde. Auf analoge Weise kommt man dann zu einer behelfsmäßigen Reliabilitätsschätzung für einen Testdurchgang, in dem der empirisch ermittelte Reliabilitätskoeffizient bei 2 Durchgängen als Basis zur Berechnung für einen auf die Hälfte reduzierten Aufgabenumfang gedeutet wurde. Eine empirische Reliabilitätsbestimmung für einen Testdurchgang war ja nicht möglich. Die weitere Analyse geht von der Annahme aus, die hohe Reliabilität von .97, die höher ausfällt als bei den früheren Studien (siehe Tabelle 1 oder 2), basiere zumindest zum Teil auf der verbesserten Version und nicht auf einer zufällig optimalen Stichprobe. Aber alle, auch die früheren Befunde, sprechen dafür, dass zumindest für die Bearbeitungsgeschwindigkeit des GU auch bei einem Testdurchgang noch eine Reliabilität von mindestens ca. $\alpha = .80$ erhalten bleibt.

Tabelle 2: GU_Testwerte bei unterschiedlicher Anzahl von Testdurchgängen (WS 2015/16)**a) Reliabilität α Bearbeitungszeit**

Anzahl der Durchgänge	empirisch ermittelt	nach Spearman-Brown
3	.97	-
2	.95	.96
1	-	.90

b) Reliabilität α Konzentrationsleistung

Anzahl der Durchgänge	empirisch ermittelt	nach Spearman-Brown
3	.89	-
2	.88	.84
1	-	.79

Die Reliabilitäten für die Konzentrationsleistung fallen erwartungsgemäß etwas niedriger aus, weil darin auch die weniger reliablen Fehler verarbeitet sind. Bei der Abwägung von Testaufwand und Reliabilität sind alle Koeffizienten in Tabelle 2 als erwägenswert einzustufen. Für Individualdiagnosen könnten bereits zwei Durchgänge genügen und in Gruppenstudien mit relativ passablen Gruppengrößen wäre aus ökonomischen Gründen selbst ein Testdurchgang in Erwägung zu ziehen.

Validität und Testlänge

Auf die als gegeben angesehene inhaltliche Validität des GU wurde bereits hingewiesen. Ein relativ einfaches empirisches Validitätsmaß überprüft die Beziehungen des Tests mit bereits bekannten Verfahren, die Ähnliches messen sollen (konvergente Validität). Bei den Validitätsstudien kamen nachfolgende Tests zum Einsatz:

ZRF_20_5: Zahlenreihenfolgetest (Jacobs 2015)

TMT_B_v: Variante des Trail Making Part B nach Reitan (1992)

B-STT: Variante des Berliner Shape Trail Test nach Riemer (2015)

BSZ: Konzentrationstest Buchstabe, Symbol, Zahl [Eigenkonstruktion]

d2c: Eine Variante des d2 nach Brickenkamp et al. (2010)

REV_c Eine Variante des Revisionstests nach Marschner, G. (1972)

KLT: eine Variante des Konzentrationsleistungstest nach Düker & Lienert (2001)

ZZ: Anzahl von 1 und 7 aus 200 einstelligen Zahlen zählen nach Manthei (2006)

Tabelle 3 fasst die Korrelationen zwischen dem GU_3 (=GU mit drei Testdurchgängen) und etlichen Konzentrationstests zusammen. Insgesamt liegen die Korrelationshöhen im erwarteten Rahmen. Tabelle 3b differenziert die Korrelationen der neuesten GU-Version aus der Studie im WS 2015/15 nach der Anzahl der Testdurchgänge des GU.

Tabelle 3 : Ergebnisse von Validitätsstudien zum GU

a: Zusammenhang zwischen GU_3 und etlichen Konzentrationstests

	ZRF_20_5	KLIT	ZZ	REV_c	d2c	TMT_B_v	B-STT	BSZ
Jacobs								
2013	.50	.44	.22					
2014	.44			-.64	.50			
WS 15/16	.45				.64	.51	.44	.61

b.) Korrelationen zwischen GU Varianten mit unterschiedlicher Anzahl von Testdurchgängen und sonstigen Konzentrationstests aus der Studie im WS 2015/16

	ZRF_20_5	d2c	TMT_B_v	B-STT	BSZ
GU_1	.41	.62	.49	.43	.57
GU_2	.43	.62	.51	.44	.60
GU_3	.45	.64	.51	.44	.61

GU_1 = erster Durchgang

GU_2 = (erster + zweiter Durchgang)/2

GU_3 = (erster+zweiter+dritter Durchgang)/3.

Wie man in Tabelle 3b sieht, gehen mit sinkender Anzahl der Testdurchgänge keine nennenswerten Validitätseinbußen einher, vermutlich deshalb, weil auch weniger Testdurchgänge eine hinreichende Zuverlässigkeit gewährten. Im Nachhinein sieht es so aus, dass hoch vergleichbare Validitätskoeffizienten ermittelt worden wären, wenn man auf die beiden letzten Testdurchgänge des GU verzichtet hätte. Für etliche Forschungsfragen dürfte eine Verkürzung des GU auf 2 Durchgänge oder einen Testdurchgang mehr Vor - als Nachteile erwarten lassen.

Orientierungswerte für studentische Leistungen

Bisherige empirische Studien wurden ausschließlich mit Studierenden durchgeführt und nur für diese können Vergleichsdaten angeführt werden. In der Version 3 findet man weder für die Bearbeitungsgeschwindigkeit noch die Konzentrationsleistung Zusammenhänge mit Geschlecht, Alter oder Abiturnotendurchschnitt.

Da als Testwert des GU die durchschnittliche Bearbeitungszeit aus den realisierten Testdurchgängen herangezogen wird, liefern diese auch gewisse Orientierungswerte für Versionen, die mehr oder weniger Durchgänge durchlaufen. Diese These setzt freilich voraus, dass während einer Testung mit mehrfachen Testdurchgängen keine nennenswerten Übungsfortschritte erzielt werden. Im Verlauf einer [der ersten] Testung mit 3 Durchgängen ließen sich jedoch signifikante Verbesserungen feststellen, die sich im Wesentlichen von Durchgang 1 nach 2 und 1 nach 3, aber nicht mehr von 2 nach 3 vollzogen. Das Ausmaß der Zeitgewinne fällt mit einer Effektstärke d von höchstens .14 aber vernachlässigbar gering aus.

Tabelle 4**a) GU-Bearbeitungszeiten** aufgeteilt in 1,2 und 3 Testdurchgänge (Studie WS 15/16)

	α	M	s	Md	Schiefte	Kurtosis	N
GU_1	.90	97.7	23.4	93.5	.66	-.45	98
GU_2	.95	96.4	23.4	89.5	.68	-.40	98
GU_3	.97	95.7	23.4	90.5	.80	-.01	98

b) GU-Konzentrationsleistung aufgeteilt in 1,2 und 3 Testdurchgänge (Studie WS 15/16)

	α	M	s	Md	Schiefte	Kurtosis	N
GU_1	.79	14.6	4.0	14.8	.0	.69	98
GU_2	.88	14.5	3.9	14.5	.0	.02	98
GU_3	.89	14.2	3.7	14.6	.05	.52	98

Tabelle 4 stellt die deskriptiven Ergebnisse des GU für unterschiedliche Anzahlen von Testdurchgängen dar. GU_1 besteht aus dem ersten Durchgang, GU_2 aus den ersten beiden und GU_3 aus allen 3 Durchgängen. Bzgl. der durchschnittlichen Bearbeitungszeiten waren bei den drei GU_Varianten zwar noch signifikante, aber von der Effektstärke her äußerst marginale Zeitunterschiede zwischen GU_1 und GU_2 ($d=.06$), GU_1 und GU_3 ($d=.09$), sowie GU_2 und GU_3 ($d=.03$) zu konstatieren. Auch hinsichtlich der Konzentrationsleistung fallen die Unterschiede, pragmatisch betrachtet, ziemlich unerheblich aus. Es sieht so aus, dass bei einer [,der ersten,] Testbearbeitung die Anzahl der Testdurchgänge von 1 bis 3 für das Testergebnis kaum eine Rolle spielt. Im Anhang findet der Leser Normen für Studierende, die auf dem GU_3 aus der Studie im WS 15/16 basieren. Wenngleich es sich bei diesen Normen nicht um eine klassische Normierung handelt – im strengen Sinne weder repräsentativ noch vom Umfang hinreichend groß – liefern Sie meiner Meinung nach eine realistische Orientierung für Studierende, weil die Studierenden der untersuchten Stichprobe (Lehramtsstudierende) ein großes Spektrum unterschiedlicher Fächer studieren.

Übungseffekte bei Testwiederholungen

Grundsätzlich bewirkten wiederholte Testungen des GU_3 im Sinne des verteilten Trainings sehr deutliche Übungsgewinne. So führten 11 Testungen des GU_3 (=33 Testdurchgänge insgesamt) im Verlauf einer Woche zu einer Zeitersparnis von 45% (Jacobs 2014). Mit zunehmender Testung entwickeln die Probanden effektive Lösungsstrategien und können die Zielzahlen mit immer weniger geistigem Aufwand identifizieren. Hierbei erzielten die zu Beginn konzentrationschwachen Studierende den höchsten Leistungszuwachs bzw. Zeitgewinn. Durch dieses umfangreiche Training reduzierte sich auch die Reliabilität beim abschließenden Nachtest. Schon bei einer einfachen Wiederholungsmessung -Vortest und nach ca. 7 bis 10 Tagen Nachtest - fand Jacobs (2014) mit dem GU_3 einen reinen Testeffekt in der Effektstärkengröße von $d=.78$, ein Übungsgewinn, wie er für viele Konzentrationstests, etwa auch den d2, keineswegs unüblich ist. Hier würde die Reduktion der Testdurchgänge den Übungseffekt vermutlich deutlich eindämmen. Höchstwahrscheinlich wäre der Testeffekt ganz zu vermeiden, wenn man im Vortest den GU und im Nachtest den UG verwenden würde.

Literatur

- Brickenkamp, R. Schmidt-Atzert, L., Liepmann, D. (2010). d2 –Revision Aufmerksamkeits- und Konzentrationstest - Manual. Hogrefe
- Düker, H., Lienert, G. A. (2001) Konzentrations-Leistungs-Test Revidierte Fassung (KLT-R) Neubearbeitung von: Lukesch, H. & Mayrhofer, S. Hogrefe Verlag, Göttingen.
- Jacobs, B. (2002a). Testanweisung: Informeller Zahlenkonzentrationstest
http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/gerade_ungerade/altversion/guanweisung.htm [5.1.2013]
- Jacobs, B. (2002b). Informeller Zahlenkonzentrationstest - Identifiziere eine zweistellige Zahl: erst gerade, dann ungerade Ziffer
http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/gerade_ungerade/altversion/gukonstruktion.htm [5.1.2013]
- Jacobs, B. (2013). Erprobung zweier Online-Konzentrationstests mit Zahlen an Studierenden des Lehramts.
<http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/konzentrationstests.html>
- Jacobs, B. (2014): Analyse von Testgütekriterien und Übungseffekten zweier Online-Konzentrationstests.
 URN: urn:nbn:de:bsz:291-psydok-49678
 URL: <http://psydok.sulb.uni-saarland.de/volltexte/2014/4967/>
- Jacobs, B. (2015a) Der Zahlenreihenfolgetest 20_5 Psychometrikon - psychologisch medizinisches Testportal - <http://psychometrikon.de/inhalt/suchen/test.php?id=c98cf246e31e3cd376cdf94ad04df78> DOI: 10.6099/1000211
- Jacobs, B. (2015b). Zusammenhänge zwischen Leistungsprognosen, Ergebniseinschätzungen, erzielten Leistungen in Konzentrationstests und subjektiver Auffassung der eigenen Konzentrationsfähigkeit. ResearchPaper zugänglich unter
 URN: <urn:nbn:de:bsz:291-psydok-52522>
 URL: <http://psydok.sulb.uni-saarland.de/volltexte/2015/5252/>
- Manthei, K. (2004). Pay - But Don't Pay Too Much: An Experimental Study on the Impact of Incentives, Department of Personnel Economics, University of Cologne, DP 04-27 K.
<http://www.whu.edu/orga/geaba/Papers/2004/Manthei.pdf> [7.7.2006]
- Marschner, G. (1972) Revisions-Test (Rev.T.). Göttingen: Hogrefe.
- Reitan, R.M. (Ed.). (1992). Trail Making Test. Tucson, AZ:Reitan Neuropsychology Laboratory. zitiert nach: Rodewald, K. et al. (2012)
- Riemer, T. G. (2015). Der Berliner Shape Trail Test – Charakterisierung und Validierung eines neuen Trail Making Tests (Doctoral Dissertation, Freie Universität Berlin).
- Westhoff, K. & Hagemeyer, C. (2005). Konzentrationsdiagnostik. Pabst, Lengerich.

Anhang

Nachfolgende Tabelle zu Normwerten für den GU_3 basiert auf der Stichprobe des WS 15/16. Die Rohwerte (Bearbeitungszeiten) wurden unter Zuhilfenahme des Online Normal Rank Transformators (<http://www.psychometrica.de/norming.html> [8.7.2015]) nach dem Ranking Algorithmus Blom in Prozenträge und T-werte umgewandelt.

Bei der Bearbeitungszeit wurden Prozenträge und T-Werte so umgepolt, dass hohe Werte hoher Konzentrationsfähigkeit entsprechen.

Ganz wichtiger Hinweis

Alle Angaben gelten nur für die erste Testbearbeitung

Tabelle 5: Normen für Studierende im GU_3 (N= 98; Daten aus WS 15/16)

a) Bearbeitungszeiten

Sekunden	Prozent-rang	T-Wert
<60	99	74
60-64	97	68
65-69	96	67
70-74	88	62
75-79	75	57
80-84	66	54
85-89	54	51
90-94	50	50
95-99	44	48
100-104	35	46
105-109	30	45
115-119	19	41
120-124	17	40
125-129	13	39
130-134	9	37
134-139	8	36
140-144	7	35
145-149	6	34
150-154	2	30
>155	1	26

Standardmessfehler: 4.1

b) Konzentrationsleistung (KL), auf ganze Zahlen gerundet
Anzahl fehleradjustierter korrekter Angaben pro Minute

KL	%Rang	T-Wert
4	1	25
7	2	30
9	7	35
10	14	39
11	21	42
12	29	44
13	36	46
14	44	49
15	55	51
16	66	54
17	77	58
18	87	61
19	93	65
20	95	67
21	97	69
22	98	71
26	99	75

Standardmessfehler: 1.2