

Analyse von Testgütekriterien und Übungseffekten zweier Online-Konzentrationstests

Abstract

Mithilfe zweier von Jacobs (2013) konstruierter Online-Kurzzeitkonzentrationstests wurde der Einfluss mehrfacher Testwiederholungen auf die Konzentrationsleistungen geprüft. Zu diesem Zweck wurden Studierende nach Zufall auf 2 Gruppen zugeteilt und bearbeiteten jeweils einen Konzentrationstest insgesamt zehnmal, während der andere Konzentrationstest als Kontrolle diente. Bei einem der Konzentrationstests führte bereits die erste Testwiederholung zu besseren Testresultaten und das zehnmahlige Testen bewirkte markante Übungsgewinne, während beim anderen Konzentrationstest die erste Testwiederholung keinen Effekt nach sich zog und die Trainingswirkung nach 10 Testwiederholungen zwar statistisch nachweisbar, aber deutlich geringer ausfiel. Die deskriptiven Ergebnisse bestätigten weitgehend die bisherigen Erfahrungen und die Reliabilitäten der beiden Onlinetests fielen zufriedenstellend aus. Ansprechende Korrelationen der Tests mit speziellen Computervarianten des d2 und des REV-T unterstützen die Validität der Onlinetests und sprechen im Verbund mit den bisherigen Erkenntnissen insgesamt dafür, dass sie im Wesentlichen Konzentration erfassen. Zumindest einer der Onlinetests erweist sich als besonders nützlich, als Messinstrument zur Erfassung von Konzentration im Rahmen von Onlinestudien Verwendung zu finden.

Einleitung und Zielsetzung.

Ziel der Studie war es, zwei kürzlich entwickelte Onlinekonzentrationstests, den GU und den ZRV_20 (Jacobs 2013) weiter auf ihre Nützlichkeit hin zu untersuchen. Zum einen sollten die bisherigen Befunde repliziert, Retestkorrelationen berechnet, weitere Hinweise für ihre Validität gefunden sowie vornehmlich der Einfluss des Trainings durch Testwiederholungen ermittelt werden. Es ist schon lange bekannt, dass die Ergebnisse von Konzentrationstests in hohem Maß davon abhängen, wie häufig diese Tests zuvor bearbeitet wurden (z.B. Westhoff (1989), Westhoff & Dewald (1990)). Die Übungsgewinne können enorm ausfallen und bei intensivem Training locker 50% der Ausgangsleistung oder noch mehr ausmachen. Dabei beziehen sich die Übungsgewinne leider nicht auf eine Steigerung der allgemeinen Konzentrationsfähigkeit, sondern lediglich auf die Leistungen des trainierten Konzentrationstests. Hier soll sich diese fehlende Transferwirkung darin zeigen, dass Probanden, welche nur einen der beiden Tests intensiv einüben, in diesem massive, im nicht trainierten Test jedoch nur marginale Vorteile erzielen. Wie hoch die potenziellen Trainingswirkungen ausfallen, hängt vermutlich auch vom speziellen Konzentrationstest ab.

Informelle Selbstversuche des Verfasser verstärkten bei diesem die Ansicht, schon in wenigen Versuchen große Fortschritte im GU erzielt zu haben, bei dem aus einer Menge zweistelliger Zahlen diejenigen herausgesucht werden müssen, deren erste Ziffer gerade und deren zweite Ziffer ungerade ist, aber mehrfache Versuche kaum einen Trainingseffekt beim ZRV_20 nach sich zogen, bei welchem die zufällig angeordneten Zahlen 1 bis 20 entsprechend ihrer Reihenfolge angeklickt werden sollten. Lassen sich die Vermutungen in einem Gruppenversuch statistisch bestätigen, so ist nachgewiesen, dass die Übungsgewinne offenbar auch vom speziellen Konzentrationsverfahren abhängen. Vom diagnostischen Standpunkt aus wären besonders solche Konzentrationstests nützlich, die nur sehr geringe Übungseffekte bewirken, weil so zumindest die Verfälschbarkeit im Hinblick auf gute Leistungen eher eingegrenzt werden könnte. Daneben wäre es vorteilhaft,

einige empirische Argumente beizusteuern, die begründeten Anlass zu der Vermutung geben, durch die Onlinetests im wesentlichen Konzentration zu messen.

Zunächst wird der Versuchsaufbau, anschließend die Ergebnisse zu einzelnen Konzentrationstests und abschließend eine Gesamtschau aller Ergebnisse dargestellt.

Versuchsdurchführung und Versuchsaufbau.

An dem Versuch nahmen bis zu 69 Studierende aus 3 Seminaren des Verfassers im WS 2013/14 teil. Das Durchschnittsalter betrug 24.5 Jahre, ca. 2/3 der Studierenden waren Frauen.

Den Studierenden wurde mitgeteilt, Versuchsaufbau und Daten der Untersuchung würden im Rahmen der Seminare für didaktische Zwecke zur Verdeutlichung etlicher Themen wie etwa Versuchsplanung, Reliabilitätsanalysen, Konzentrationstestung usw. benötigt. Da die gesamte Untersuchung online abgewickelt wurde, bekamen die Studierenden für einzelne Versuchsabschnitte jeweils ein Zeitfenster gewährt, um hinreichende Gelegenheit zu bieten, zu Hause in Ruhe Konzentrationstests am eigenen Computer zu bearbeiten. Tabelle 1 gibt eine Übersicht zu den Erhebungszeitpunkten, dem Erhebungszweck und den eingesetzten Konzentrationstests.

Tabelle 1: Übersicht zur Erhebung

Datum	14.11.2013- 17.11.2013	18.11.2013- 23.11.2013	25.11.2013- 27.11.2013
Zweck	Vortest	Training	Nachtest
Testverfahren	GU ZRF_20 d2c	10 mal GU oder 10 mal ZRF_20	GU ZRF_20 d2c

Zunächst bearbeiteten alle Probanden als Vortest 3 Online-Konzentrationstests. GU und ZRF_20 wurden von Jacobs 2013 übernommen. Der d2c ist eine spezielle Computervariante und basiert auf dem bekannten Aufmerksamkeits- und Konzentrationstest d2 (Brickenkamp et al. 2010). Dieselben Konzentrationstests waren dann von allen Probanden auch im Nachtest zu bearbeiten.

Tabelle 2 verdeutlicht das Vorgehen zur Prüfung der Trainingseffekte. Vor der Trainingsphase wurden alle für die Studie verfügbaren Probanden nach Zufall auf 2 Gruppen, A und B, zugeteilt worden. Die Studierenden unter A sollten in der Trainingsphase innerhalb einer Woche 10 mal den Konzentrationstest GU bearbeiten, während die Probanden unter B 10-mal den Konzentrationstest ZRF_20 zu absolvieren hatten. Zur Prüfung der Trainingseffektivität des GU dient Gruppe A als Experimentalgruppe (EG) und B als Kontrollgruppe (KG). Umgekehrt fungiert B als EG und A als KG zur Überprüfung der Wirksamkeit des Trainings beim ZRF_20.

Tabelle 2: Versuchsplan zur Prüfung von Trainingseffekten

	Vortest	Training	Nachtest
A	GU	10 * GU	GU
B	GU		GU
A	ZRF_20		ZRF_20
B	ZRF_20	10 * ZRF_20	ZRF_20

Ein spezielles Training des d2c war nicht vorgesehen. Die Studierenden bearbeiteten den d2c lediglich im Vor- und Nachtest, um so die Retestreliabilität des d2c zu überprüfen und eine weitere Kontrolle möglicher Mittelwertsveränderungen ohne Training zur Verfügung zu stellen. Außerdem diente der Test als Validitätskriterium für die beiden übrigen Tests.

Bei allen eingesetzten Konzentrationstests folgte nach der Bearbeitung eine Rückmeldung zum erzielten Ergebnis. Diese beschränkte sich bei d2c auf die Mitteilung der Bearbeitungszeit und der Fehler, fiel beim GU und ZRF_20 aber noch ausführlicher aus. Dort bekamen die Studierenden Vergleichsdaten (M und s) und ihre Rangposition auf der Basis einer für die Testpersonen repräsentativen Stichprobe von ca. 100 Studierenden dargeboten. Diese Rückmeldungen sollten die Motivation beim Training hoch halten und haben diese Funktion vermutlich auch ausgeübt, da Konzentrationsübungen in der Regel zu Leistungsverbesserungen führen, die der Testperson Erfolge ermitteln und die Erfolgserwartung erhöhen.

Für die Trainingsphase erhielten die Probanden einen Plan zur Selbstkontrolle. Dieser sah vor, von montags bis freitags zweimal täglich den Konzentrationstest einzuüben. Als Ausweichtermin war der Samstag gedacht, falls das Training an einem der vorherigen Tage nicht durchgeführt werden konnte. Durch die zeitliche Streckung der Trainingsphase sollte eine verteilte Übung angeregt werden, weil diese einen nachhaltigeren Übungseffekt versprach. Während die meisten Studierenden dem vorgegebenen Plan angemessen folgten, hielten sich einige nicht minutiös an diese Vorgaben, was teilweise auch zu Übungsausfällen führte, auf die bei der Auswertung näher eingegangen wird.

Die Haupthypothese lautet, durch die Einübung eines speziellen Konzentrationstests verbessere sich die Leistung in diesem Test, im nicht trainierten Konzentrationstest hingegen nicht bzw. nicht über das Maß hinaus, was bei einer Wiederholungsmessung von Konzentrationstests ohnehin erwartet werden muss. Für die strenge statistische Prüfung bedeutet dies die Erwartung einer größeren Nachtest-Vortest-Differenz für die Gruppe mit eingeübtem Konzentrationstest (EG) im Vergleich zur Gruppe ohne das entsprechende Training (KG). Darüber hinaus wird als schwächere Forderung innerhalb jeder Gruppe die bedeutsamste Konzentrationsverbesserung im eintrainierten Konzentrationstest erwartet, was allerdings die umstrittene Annahme voraussetzt, alle Konzentrationstests seien in annähernd vergleichbarem Maße trainierbar.

Bevor die Daten ausgewertet werden konnten, war eine gründliche Inspektion erforderlich, da mit Fehlern gerechnet werden musste, welche die Ergebnisse massiv beeinflussen könnten. Während Auslassungsfehler (= Zielitems werden nicht angeklickt) noch zu einem gewissen Ausmaß akzeptabel waren, aber ab einem bestimmten Niveau nicht mehr tragbar erschienen, wurden obligatorisch alle Daten des Konzentrationstests ausgeschlossen, sofern eine Testperson ungewöhnlich häufig die falschen Items ankreuzte. Glücklicherweise kam letzteres sehr selten vor. Ein weiteres Problem liefern Ausreißer bei der Bearbeitungszeit. Teilweise wurden die entsprechenden Testwerte gar nicht gewertet, teilweise auf einen sehr hohen Zeitwert herunter getrimmt. Manche Probanden hielten sich stellenweise nicht strikt an die Anweisungen und übersprangen so z.B. einen Testdurchgang. Für bestimmte Analysen (etwa Reliabilität) war der Datensatz nicht mehr nutzbar, für andere Analysen konnten die verbliebenen Werte aber noch verwendet werden. Insgesamt bewirkten die Fehler gewisse Datenausfälle, was dazu führt, dass

die nutzbaren Daten für verschiedene Vergleiche auf einer unterschiedlichen Anzahl von Probanden basieren.

GU: [Kreuze alle Zahlen an, deren erste Ziffer gerade und deren zweite Ziffer ungerade ist]

Im GU-Konzentrationstest muss die Testperson in insgesamt 3 Durchgängen jeweils aus 100 zweistelligen Zahlen die 25 Zielzahlen anklicken, deren erste Ziffer gerade und deren zweite Ziffer ungerade ist (siehe Anhang 1 bzw. Jacobs 2013). Als Konzentrationsmaße wurden hier erhoben:

- **GU-genauigkeit:** Prozentsatz der korrekt identifizierten Zielitems [bei Jacobs 2013 GU-leistung genannt]
- **GU-zeit:** durchschnittliche Bearbeitungszeit pro Durchgang in Sekunden
- **GU-leistung:** GU-genauigkeit/GU-zeit

Als entscheidendes Konzentrationsmaß für den GU-Konzentrationstest gilt das Arbeitstempo bzw. die benötigte Zeit. Sie wird als durchschnittliche Bearbeitungszeit aus 3 Durchgängen in Sekunden erfasst. Dieses Zeitmaß wäre dann als hinreichend valide einzuschätzen, wenn es auf einer absolut fehlerfreien Bearbeitung basierte. Da die GU-genauigkeit den Prozentsatz der identifizierten Zielitems angibt, entspricht die Differenz $100 - \text{GU-genauigkeit}$ dem Prozentsatz der Auslassungen. Das Genauigkeitsmaß erfasst also zugleich den entscheidenden Fehler, zumal angeklickte Distraktoren äußerst selten vorkamen und deshalb unberücksichtigt bleiben. Wie eine Überprüfung durch Varianzanalysen mit Messwiederholungen ergab, ließen sich bei der ersten GU-Testung keine Verbesserungen bzgl. Genauigkeit und Bearbeitungsgeschwindigkeit im Verlauf der 3 Durchgänge feststellen. D.h. Im Verlauf der ersten Testung sind keine Übungseffekte nachzuweisen.

Einer Analyse von Jacobs (2013) zufolge erbrachte die Genauigkeit (dort GU-leistung genannt =Prozentsatz der korrekten Lösungen) allein allerdings keine befriedigende Reliabilität. Viele Genauigkeitsfehler würden allerdings erhebliche Interpretationsprobleme verursachen, wenn dadurch Bearbeitungszeit eingespart werden würde. Das scheint jedoch eher nicht der Fall zu sein, da der Zusammenhang zwischen Genauigkeit und Bearbeitungszeit von $r = .13$ ($N=63$) im Vortest ähnlich wie Jacobs 2013 insignifikant ausfiel. Als zusätzliches Konzentrationsmaß wurde die GU-leistung als Quotient aus Prozentsatz korrekter Lösungen geteilt durch Bearbeitungszeit gebildet. Dieses Maß liefert allerdings keinen deutlichen Mehrwert gegenüber der einfachen Bearbeitungszeit, da beide Variablen im Vortest $r = -.84$ miteinander korrelieren. Um unnötige Redundanz zu vermeiden, werden hauptsächlich die Ergebnisse zur Bearbeitungszeit mitgeteilt und nur vereinzelt auf die GU-leistung eingegangen.

Zunächst wurden die Daten einer augenscheinlichen Fehleranalyse unterzogen. Hierbei mussten einige Daten von 2 Probanden ausgeschlossen werden, da diese ganz deutliche Auslassungen oder massive Falschlösungen hinterließen. 3 Probanden erzielten sehr extreme Zeiten beim Vortest, die im Tukey-Boxplot als eindeutige Ausreißer erkennbar waren und auf 150 Sekunden (2 s über dem Mittelwert) getrimmt wurden.

Eine deskriptive Analyse der Untersuchungsstichprobe ergab für die untersuchte Gesamtstichprobe hinsichtlich Mittelwerte, Streuungen und Reliabilität für die Bearbeitungszeiten hoch vergleichbare Ergebnisse zu den Befunden von Jacobs 2013. Tabelle 3 zeigt die Ergebnisse etlicher Reliabilitätsanalysen.

Tabelle 3: GU: Reliabilität α und Retest (ca. 10-12 Tage)

			Genau- igkeit G	benötigte Zeit Z	Leistung G/Z	N
Vortest	alle	α	.77	.90	.92	60
Nachttest	kein Training	α	.73	.92	.92	32
	Training	α	.41	.95	.95	32
Retest	kein Training	r_{tt}	.54	.83	.84	32
	Training	r_{tt}	.07	.54	.46	30

Die beste Konsistenzschätzung liefert der Vortest, da dort noch keinerlei Erfahrung mit dem GU-Konzentrationstest vorlag und 60 Probanden vollständige und akzeptable Daten vorlegten. Die Koeffizienten für die Bearbeitungszeiten um $\alpha=.90$ entsprechen den bisherigen Befunden von Jacobs (2013) während die Konsistenz der Genauigkeit hier etwas höher ausfällt.

Als beste Schätzung für die bisher unbekannt Retestkorrelation bietet sich das Ergebnis der Kontrollgruppe an, weil die Übungsdurchgänge der Experimentalgruppe zu einer deutlichen Streuungsreduktion der Bearbeitungszeiten führten. (Streuung im Nachttest mit $s = 8,8$ deutlich niedriger ist als im Vortest mit $s=21,6$). Hiermit kann der Befund von Westhoff & Dewald 1990 bestätigt werden, demzufolge Übung die Retestreliabilität mindert. Die bzgl. der benötigten Zeit ziemlich schwache Korrelation zwischen Vor - und Nachttest bei der EG lässt gewisse Zweifel an der Stabilität der Rangreihen aufkommen und vermuten, die Personen würden in unterschiedlichem Maße von der Übung profitieren. Diese These gilt zumindest für die ersten beiden Testwiederholungen. Spätestens nach der dritten Übung sehen die Zusammenhänge mit dem Nachttests aber deutlich besser aus und erreichen ab diesem Zeitpunkt Werte um .80 (siehe Anhang 3a), bei der GU-leistung bereits ab der zweiten Testung. Fasst man alle Testdurchgänge der EG als Items eines umfassenden Konzentrationstest GU, dann beträgt das α für die Genauigkeit .57, das für die Bearbeitungszeit .97 und für die GU-leistung .98.

Vergleich Training vs. kein Training:

Einige wenige Probanden in EG und KG ließen höchstwahrscheinlich aus Unachtsamkeit einen von 3 Durchgängen des GU unbeantwortet. Um eine möglichst hohe Probandenanzahl für den Vergleich zwischen EG und KG zu ermöglichen, wurden deren Vor - oder Nachttests auf der Basis von 2 statt 3 Durchgängen berechnet. Die dadurch bedingte Reliabilitätseinbuße wirkt sich nicht bedeutsam auf den Gruppenvergleich aus.

30 Probanden der Trainingsgruppe lieferten verwertbare Daten in Vor- und Nachttest. Davon bearbeiteten 22 Testpersonen anweisungskonform 10 GU-Übungen, 7 Testpersonen nahmen mindestens 7 und höchstens 9 GU-Tests in Angriff und eine Person 5 Übungen. Testpersonen, die weniger als 5 Übungen absolvierten, wurden aus der Trainingsgruppe ausgeschlossen.

Bei der **Genauigkeit** schwanken die Prozentsätze korrekter Lösungen bei den insgesamt 12 Messungen der EG zwischen 89% und 99%. Wie aus Tabelle 4 hervorgeht, vollzieht sich der entscheidende Genauigkeitszuwachs von der ersten zur zweiten Messung, sowohl bei der EG ($t(29)=3,23$ $p=0.003$), wie bei der KG ($t(31)=3,22$; $p=0.003$). Aus diesem Grunde schneidet die EG bei der Genauigkeit im Nachttest auch nicht besser ab als die KG, bzw. präziser formuliert, die

Wechselwirkung zwischen Zeitpunkten (Vor-, Nachtest) und experimenteller Bedingung (EG, KG) bleibt bei der Genauigkeit insignifikant ($F(1,60) = 1,7$; $p=0.2$).

Tabelle 4: Genauigkeiten des GU für alle Testzeitpunkte

		Vortest	1	2	3	4	5	6	7	8	9	10	Nachtest
EG	M	89	95	96	96	97	96	96	97	96	99	98	98
	s	10	6	4	3	3	6	6	4	6	2	2	2
KG	M	89											94
	s	11											7

Da die EG-Probanden im Nachtestmittelwert aber 98% korrekte Angaben machten, werden auch deren schwache Konsistenz sowie die fehlende Retestrelia-
 abilität verständlich.

Die für die Hypothesenprüfung entscheidenden Analysen werden mit der Bearbeitungszeit GU-zeit als abhängige Variable durchgeführt. Wegen der sehr hohen Korrelation zwischen GU-zeit mit der GU-leistung sind bei der GU-leistung keine wesentlich anderen Ergebnisse zu erwarten.

Zunächst ergab ein Vergleich der Vortestwerte von EG und KG keine signifikanten Unterschiede bzgl. Genauigkeit, Bearbeitungszeit und GU-leistung (p nach t -test jeweils größer gleich .21) und bestätigte so die Effizienz der Randomisierung. Wie aus Tabelle 5a hervorgeht, erbrachte die zweifaktorielle Varianzanalyse mit dem between faktor Experimentelle Bedingung (Training, kein Training) und dem repeated measurement factor Zeitpunkte (Vor- bzw. Nachtest) signifikante Effekte auf beiden Faktoren sowie der Wechselwirkung.

Tabelle 5a: Ergebnis der Varianzanalyse für die Bearbeitungszeit des GU

Faktor	F(1,60)	p
Zeitpunkt (Z)	188,0	.000
Exp. Bedingung (E)	20,0	.000
Interaktion Z x E	32,6	.000

Tabelle 5b: Mittelwerte und Streuungen der GU-Zeiten (sec)

		Vortest		Nachtest	
N		M	s	M	s
EG	30	93,5	20,8	53,3	9,5
KG	32	101,9	25,1	84,3	15,8

Tabelle 5b und Abbildung 1 verdeutlichen die Effekte und veranschaulichen insbesondere den klaren Trainingseffekt in Form der Interaktion, welche den deutlich markanteren Übungsgewinn durch die wiederholten Testungen wieder-
 spiegelt. Die Vortest korrigierte Effektstärke für den Trainingseffekt bzgl. der Bearbeitungsgeschwindigkeit beträgt $d=1.99$ (für die GU-leistung $d=1.96$).

Abbildung 1

GU für Trainings- und Kontrollgruppe

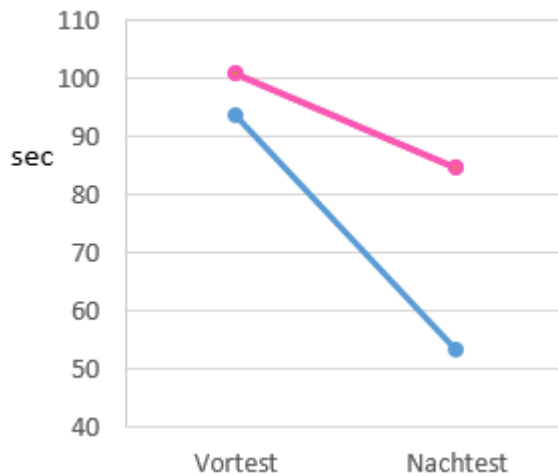
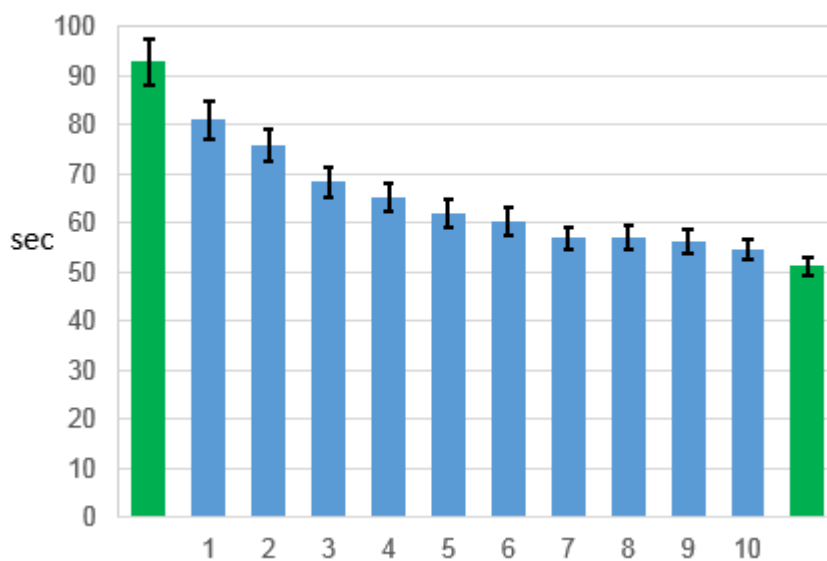


Abbildung 2 zeigt die Messwerte für die 22 Probanden der Experimentalgruppe mit vollständigen Daten zu allen Zeitpunkten. Bereits die erste Testwiederholung erbrachte einen signifikanten Übungsgewinn ($t(21)=2,9$, $p<0,01$), der einer Effektstärke von $d=.59$ entspricht. Der Unterschied zwischen Vor- und Nachtest fällt mit einer Effektstärke von $d=2.53$ noch etwas höher aus als der oben angestrengte Vergleich mit der Kontrollgruppe und unter Einbezug solcher Studierender in die EG, welche ihr Übungssoll nicht ganz erfüllt hatten. Auch die Kontrollgruppe erzielte bei der zweiten Messung (=Nachtest 10 bis 12 Tage nach Vortest) einen ganz klaren Testeffekt ($t(31)=6,31$ $p<0,001$; Effektstärke $d=0.79$; siehe Abbildung 1)

Abbildung 2

GU: 10 Übungen zwischen Vor- und Nachtest



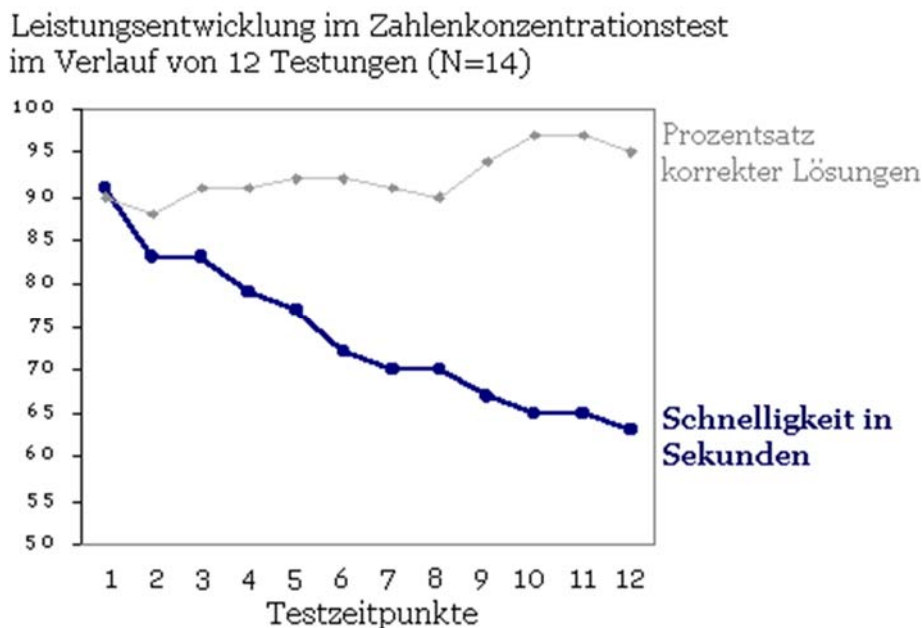
Anmerkung: die Fehlerbalken entsprechen dem Standardfehler des Mittels.

Gemessen am Vortest benötigte die EG im Nachtest nur 55% der für den Vortest investierten Bearbeitungszeit, ein immenser Übungsgewinn von 45% Zeitersparnis. Bei Westhoff & Dewald (1990) fielen die Vorteile mit 47% und 62 % noch etwas höher aus. Augenscheinlich sieht es so aus, dass die ersten 5 Testungen deutliche Übungsvorteile nach sich ziehen, dann die Gewinne geringer ausfallen, es allerdings unklar bleibt, ob das Ende des Leistungsniveau schon erreicht ist. Wenngleich nicht jeder Test einen signifikanten Konzentrationsgewinn gegenüber seinem vorherigen Übungstest erzielte, fällt der Nachtest der EG signifikant besser aus als alle vorangegangenen 11 Testungen (Mehrfachvergleiche nach Bonferroni, $\alpha = 5\%$ einseitig). Die Streuung der Bearbeitungszeit nimmt im Verlauf der Übung mit wenigen Ausnahmen kontinuierlich ab. [siehe analog die Fehlerbalken in Abbildung 2], ein Ergebnis, was den Befunden von Westhoff & Dewald (1990) widerspricht, die mit zunehmender Geübtheit von höheren Standardabweichungen in den Testleistungen berichten

Testpersonen der EG, die zu Beginn hohe Bearbeitungszeiten und schwache Genauigkeiten aufwiesen, profitierten am meisten von den 10 Testwiederholungen (r_s Vortest, (Nachtest-Vortest) = .81 für die Bearbeitungszeit und -.95 für die Genauigkeit (N=22; nur Probanden mit vollständigen 10 Übungstests). Aber auch die Konzentrationsverbesserungen der Kontrollgruppe gehen vornehmlich auf Personen mit hohen Vortestbearbeitungszeiten bzw. geringen Genauigkeitswerten zurück (r_s V,(N-V) = .78 bzw. -.56). Die entsprechenden Zusammenhänge zwischen geringem Ausgangsniveau zu hohem Konzentrationszugewinn fallen allerdings geringer aus, wenn statt der Bearbeitungszeit die GU_leistung herangezogen wird. Bei der GU_leistung der EG ist der Übungsgewinn sogar unabhängig vom Ausgangsniveau. Insgesamt widersprechen die Ergebnisse aber deutlich der Vermutung eines Matthäus-Effektes von Westhoff und Hagemeister (2005, S. 36) "Personen mit besseren Anfangswerten verbessern den Wert für die Bearbeitung des gesamten Tests in der Regel durch die Wiederholung mehr als Personen mit schlechteren Anfangswerten". Es gilt vielmehr das Ausgangswertgesetz: "Es ist zu erwarten, dass die Reaktionszeit durch Übung umso mehr abnimmt, je länger sie bei der ersten Testung war". (Hagemeister,2005)

Insgesamt bestätigen die Befunde mit verblüffender Ähnlichkeit die Ergebnisse einer informellen Studie von Jacobs (2002) mit einer noch deutlich kürzeren Vorgängervariante des GU (siehe Abbildung 3). Hierbei hatten 14 Studierende den Auftrag, zu Hause über 3 Tage verteilt insgesamt 12 mal jeweils nur einen GUDurchgang der Vorgängerversion zu absolvieren und -im Gegensatz zur vorliegenden Studie, welche eine automatische Datenspeicherung vornahm- selbst Ihre Ergebnisse zu notieren, die dann als Daten verwendet wurden.

Abbildung 3: Ergebnisse von Jacobs (2002) [Zahlenkonzentrationstest=GU]



http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/gerade_ungerade/altversion/kzahlenergebnis.gif
 entnommen mit freundlicher Genehmigung des Autors aus:
http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/gerade_ungerade/altversion/gukonstruktion.htm

ZRF 20 [Zahlenreihenfolge: die Zahlen 1 bis 20 in die richtige Reihenfolge bringen]

Beim Zahlenreihenfolge-Test ZRF_20 hat der Proband die Aufgabe, insgesamt 5 mal 20 ungeordnete Zahlen in der korrekten Reihenfolge anzuklicken (siehe Anhang 2 bzw. Jacobs 2013).

Als Konzentrationsmaße werden erhoben.

- **ZRF_20_zeit**: Durchschnittliche Bearbeitungszeit der korrekt gelösten Aufgaben in Sekunden
- **ZRF_20_genauigkeit**: Prozentsatz der korrekten Aufgaben (bzw. Durchgänge), der theoretisch nur die Werte 20,40,60,80 oder 100 annehmen kann.

2 EG-Probanden nahmen nicht am Nachtest teil, absolvierten aber viele Trainingssitzungen, teilweise mehr als 10. Bei diesen Probanden wurde der letzte, höchstens aber die 12. Testung als Nachtest gewertet. 7 EG-Testpersonen absolvierten weniger als die erforderlichen 10 Trainingstests, wobei jedoch 6 mindestens 8, und eine 7 Übungstests durchführten. Die entsprechenden Studierenden einbezogen, umfasste die Trainingsgruppe so insgesamt 31 Studierende, wovon 24 Testpersonen vollständige Daten abliefern.

Eine Analyse der deskriptiven Ergebnisse des ZRF_20 im Vortest ergab, dass die Untersuchungsstichprobe sowohl in der Genauigkeit ($M=89\%$, $s=17,5$) wie bzgl. der Bearbeitungszeit ($M=19,9$, $s=3,7$) recht gut vergleichbare Ergebnisse zur Studie von Jacobs 2013, Tabelle 2 erzielte. Auch die in Tabelle 6 dargestellten Reliabilitätsanalysen entsprechen, soweit möglich den bisherigen Befunden. Lediglich die Reliabilität der Genauigkeit fällt hier etwas höher aus, ist aber immer noch völlig unzureichend. Als entscheidender Konzentrationswert wird somit lediglich die benötigte Zeit angesehen. Da die ZRF_20_zeit die durchschnittliche Bearbeitungszeit nur für die korrekt gelösten Aufgaben erfasst, ergibt sich kein Validitätsproblem, Zeitgewinne auf Kosten von Ungenauigkeiten herauszuschinden. Auch die Annahme, ein höheres Risiko einzugehen, um gute Zeiten herauszuholen, erscheint unbegründet, da Bearbeitungszeit und Genauigkeit mit $r = -0.08$ unkorreliert sind.

Tabelle 6: ZRF_20: Reliabilität α und Retest: ca. 10-12 Tage:

			Genauigkeit	benötigte Zeit	N
Vortest	alle Probanden	α	.55	.87	≤ 67
Nachtest	kein Training	α	.30	.90	≤ 30
	Training	α	.47	.87	≤ 22
Retest	kein Training	Retest	.33	.78	≤ 31
	Training	Retest	.48	.67	≤ 30

Retestreliaibilitäten waren bisher unbekannt. Die beste Schätzung liefert vermutlich die Kontrollgruppe, da diese in der Zwischenzeit keine weiteren Übungen durchführte. Deren Retestreliaibilität kommt mit $r_{tt} = .78$ nicht ganz an die erwünschte Stabilität eines Konzentrationstest von mindestens $r_{tt} = .80$ heran. Die Retestreliaibilität in der EG fällt zwar numerisch geringer aus. Allerdings beträgt die Differenz KG-EG weitaus weniger als beim GU. Fasst man EG und KG zusammen und korreliert jeweils die erste mit der zweiten Testung, so ergibt sich eine Retestkorrelation von $.77$ ($N=62$).

Bei der ZRF_20 gehen nur korrekt gelöste Testdurchgänge in die Zeitrechnung ein. Die Konsistenzangabe zur ZRF_20_zeit im Vortest von $.87$ bezieht sich daher nur auf die 45 Studierenden, die alle 5 Testdurchgänge korrekt beantwortet hatten. Jacobs 2013 hat für Testdurchgänge mit weniger korrekten Bearbeitungen Schätzungen auf der Basis der Spearman Brown Formel angegeben. Hier soll nun ein empirisches Verfahren gewählt werden, dessen Ergebnisse in Tabelle 7 dargestellt werden.

Um die Reliaibilität der Zeit bei 4 korrekten Lösungen zu berechnen, wurde der letzte Zeitwert bei Studierenden mit 5 korrekten Lösungen gelöscht. Auf diese Weise konnte mit 53 Personen eine Reliaibilitätsanalyse auf der Basis von 4 Lösungen durchgeführt werden, wobei sich ein Alpha von $.88$ ergab. Analog wurde vorgegangen, um schließlich die Reliaibilität für 3 und 2 korrekte Lösungen zu ermitteln, die dann bei $.83$ bzw. $.78$ lagen.

Tabelle 7: α für den Vortest des ZRF_20_zeit bei unterschiedlicher Anzahl korrekter Durchgänge

korrekte Durchgänge	in %	α	N
	100	.87	45
4	80	.88	53
3	60	.83	55
2	40	.78	66

Für Gruppenanalysen reichen die Reliaibilitäten aus, auch wenn etliche Probanden weniger als 5 korrekte Durchgänge erzielten. Um eine hinreichend hohe Reliaibilität für eine Individualdiagnostik zu ermöglichen würde es sich anbieten, solange zu testen, bis 5 korrekte Durchgänge abgeliefert werden.

Im Anhang 3b findet man eine Interkorrelationsmatrix der 12 Testdurchgänge (10 Übungen sowie Vor- und Nachtest) und kann sich so einen Überblick der Retestkorrelationen mit unterschiedlichem Time lag verschaffen. Insgesamt bleiben die Rangunterschiede zwischen den Probanden im Übungsverlauf bestehen. Die 66 Koeffizienten schwanken zwischen $.41$ und $.90$. Fasst man alle 12 Tests der Trainingsgruppe mit vollständigen Daten zu einem umfassenden ZRF_20 zusammen, d.h. betrachtet man jeden ZRF_20 Normaltest als Item, so erzielt man ein Alpha von $.96$ ($N=24$). Aber ein analoges Vorgehen bei der Genauigkeit erreicht selbst auf diese Weise mit einem Alpha von $.60$ immer noch keinen akzeptablen Wert.

Die Genauigkeitsmittelwerte der insgesamt 12 Konzentrationmessungen bei der EG mit vollständigen 10 Übungstests schwankten ohne erkennbaren Trend zwischen 86% und 96% korrekter Durchgänge.

Tabelle 8: Genauigkeitsmittelwerte aller Testungen in EG und KG

		Vortest	1	2	3	4	5	6	7	8	9	10	Nachtest
EG	M	90	91	93	96	88	94	93	92	92	89	87	92
	s	17	20	15	13	18	11	11	14	17	13	16	12
KG	M	91											96
	s	15											10

Wie man aus der Tabelle 8 erkennt, sind für EG und KG keine deutlichen Verbesserungen vom Vor- zum Nachtest zu erkennen. Die zweifaktorielle VA mit den Faktoren Zeitpunkte (Vortest, Nachtest) und experimenteller Bedingung (EG vollständig, KG) ergab denn auch keinen signifikanten Effekt des Zeitpunkts. Der insignifikante Interaktionseffekt ($F(1,52)=0,57$ $p=.45$) belegt ähnliche Genauigkeitsentwicklungen vom Vor- zum Nachtest für EG und KG. Die insgesamt geringen Genauigkeitsunterschiede im Verlauf der Testungen sind am ehesten als Zufallseffekte zu deuten.

Wie durch die Randomisierung erwartet und bei der Genauigkeit bereits belegt, gibt es auch keine signifikanten Unterschiede zwischen EG und KG im Vortest der ZRV_20_zeit. ($t(59)=.665$; ns), was auf hoch vergleichbare Konzentrationsfähigkeiten beider Gruppen schließen lässt. Um die Wirkung des Trainings auf die Konzentrationstests umfassend zu testen, wurde eine zweifaktorielle Varianzanalyse mit dem within factor Zeitpunkte Z (Vortest, Nachtest) und dem between factor experimentelle Bedingungen E (EG=Training, KG=kein Training) durchgeführt. Da Übungen die Leistung in einem speziellen Konzentrationstest erhöhen sollten und die Bearbeitungszeit das relevante Konzentrationsmaß des ZRV_20 darstellt, war eine signifikante Interaktion erwartet worden, die sich, wie aus Tabelle 9a hervorgeht, einseitig getestet, auch bestätigte.

Tabelle 9a: Ergebnis der Varianzanalyse zum ZRF_20 Zeit.

Faktor	F(1,60)	P
Zeitpunkt (Z)	14,4	.001
Exp. Bedingung (E)	0,05	.82
Interaktion Z x E	6,1	.009

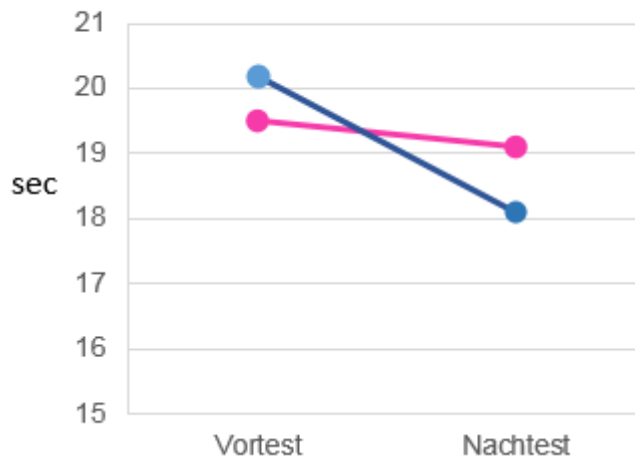
Tabelle 9b: Mittelwerte und Streuungen im ZRF_20_zeit (Sekunden)

	N	Vortest		Nachtest	
		M	s	M	s
EG	31	20,2	4,0	18,1	2,7
KG	30	19,5	3,5	19,1	3,4

Tabelle 9b und Abbildung 4 verdeutlichen eine etwas günstigere Leistungsentwicklung der EG gegenüber der KG und erneut keine Zunahme der Leistungsstreuung bei der Trainingsgruppe.

Abbildung 4: Interaktion zwischen exp. Bedingung und Zeitpunkt

ZRF_20 für Trainings- und Kontrollgruppe

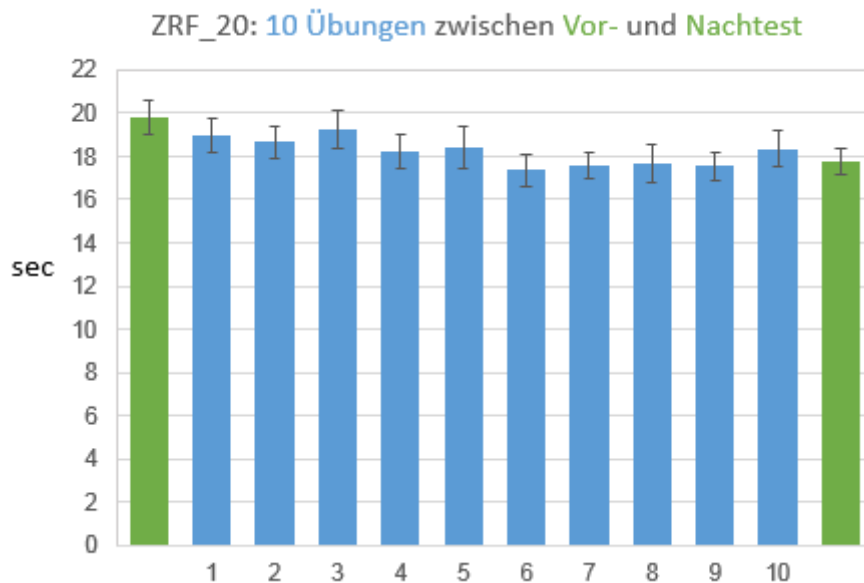


Anmerkung: Trainingsgruppe EG: mindestens 7 Übungstests; KG=keine Übungstest mit ZRF_20

Die Effektstärke mit Korrektur der Vortestwerte für den Interaktionseffekt beträgt $d = .50$. Die Ergebnisse fallen hoch vergleichbar aus, wenn man als EG nur diejenigen 24 Testpersonen einbezieht, welche die Übungen anweisungskonform zehnmal durchgeführt hatten. Die Effektstärke der relevanten Wechselwirkung zwischen den experimentellen Bedingungen und den Zeitpunkten beträgt dann $d = .49$.

Eine gesonderte Analyse mit der KG ergab, dass sich die Mittelwerte von Vor- und Nachtests nicht signifikant unterscheiden. ($t(29) = 1,13$ $pz = 0,27$). Ebenso findet man keine signifikanten Unterschiede zwischen den Mittelwerten von Vortest und erstem Übungstest bei der EG ($t(29) = 1,3$ $pz = 0,19$). Beide Ergebnisse deuten darauf hin, eine einfache Wiederholungsmessung ziehe keine deutlich besseren Ergebnisse nach sich als die erste Messung. Die insignifikante Veränderung der KG von ca. 2 % deutet klar in die Richtung, dass deren erfolgreiches Training des GU keinen Transfer auf die Leistung im ZRF_20 bewirkte, obgleich der Versuchsplan für eine ganz strenge Prüfung dieser Hypothese nicht ausreicht.

Abbildung 5 zeigt die Messwerte für die 24 Probanden der Experimentalgruppe mit vollständigen Daten zu allen Zeitpunkten.

Abbildung 5:

Anmerkung: Die Fehlerbalken repräsentieren den Standardfehler des Mittels

Augenscheinlich sieht es so aus, dass man wenigstens 4-mal den Test durchführen muss bis eine erkennbare Übungswirkung erzielt wird. Die gesamte Trainingsserie führt letztlich zu einem statistisch zuverlässigen Übungsgewinn im Nachtest. Die zwölfte Messung erbrachte gegenüber der ersten Messung einen signifikanten Zeitvorteil ($t(23)=3,81$; $pz=0,001$) in einer Größenordnung von $d=.60$. Wie oben bereits aufgezeigt, führten 10 Übungstests zwischen Vor- und Nachtest im Vergleich zur Kontrollgruppe zu einem Vorteil von $d=.49$. Vom Training profitieren vornehmlich die Probanden mit geringeren Ausgangskonzentrationswerten (=höheren Bearbeitungszeiten), denn die Korrelation zwischen Vortest und der Differenz zwischen Nachtest (N) und Vortest (V) beträgt $r_{s_{V, (N-V)}}=.55$ ($N=24$, $pz=0.005$). Wenngleich damit ein Übungseffekt beim ZRF_20 durch 10 Testwiederholungen nachgewiesen ist, so überrascht doch das für einen Konzentrationstest vergleichsweise geringe Ausmaß des Trainingseffektes, was gemessen an der Ausgangsleistung ca. 10% Zeitersparnis erbrachte. Diese geringe Zeitersparnis entspricht in etwa einer Vortestung im d2c und fällt kleiner aus als der Effekt einer einzigen Testung im GU.

Der d2 Aufmerksamkeits- und Konzentrationstest in einer Computerausführung

Auf der Basis des Manuals des Test **d2** – Revision Aufmerksamkeits- und Konzentrationstests (Brickenkamp, Schmidt-Atzert, Liepmann, 2010) wurde ein Kurzzeittest als Computervariante konstruiert, -im folgenden d2c-Test genannt-.

Diese übernimmt die Erklärung zur Testbearbeitung weitgehend original getreu und passt sie lediglich auf die besondere Kennzeichnung durch den Computer an. Vor der Testung kann die Testperson verschiedene Zeichen ausprobieren und auf ihre Korrektheit hin prüfen. Testorganisation und Testablauf unterscheiden sich jedoch deutlich von der Papierfassung.

Die Computervariante umfasst insgesamt nur 250 Zeichen, die in 2 Testdurchgängen mit jeweils 5 Zeilen von je 25 Zeichen ohne jede Zeitbegrenzung zur Bearbei-

tung vorgelegt werden. Hierbei entsprechen pro Durchgang 53 Zeichen den Zieli-temen und 72 Zeichen den Distraktoren. Die Auswahl der Ziel- und Distraktoreni-temen orientiert sich in etwa am Original. Allerdings wird die Positionierung der Zeichen stets nach Zufall vorgenommen. Abbildung 6 zeigt somit nur eine der vielen möglichen Anordnungen.

Abbildung 6: Bildschirmkopie eines möglichen d2c-Durchgangs

Durchgang Nr.1

```

" d d p d p d d d d d p d d d d p d p p d p d p d
  " " |   |   |   |   |   |   |   |   |   |   |   |   |
| p d p p p d d d d p d p p d d p d p p d d p p d
" " " |   |   |   |   |   |   |   |   |   |   |   |
d d p d d p d d d p d p d d p d d d d d d p d d d
| " | |   |   |   |   |   |   |   |   |   |   |   |
p d d d d d p p p d d d d p d d d d d d p d d d p
" | | |   |   |   |   |   |   |   |   |   |   |   |
" p p p p d d d d p p p d d p d d p d p d p d d d p
| | | | " " " | | |   |   |   |   |   |   |   |   |

```

bestätigen!

Die Testperson hat die Aufgabe, **so schnell wie möglich, aber dennoch korrekt** alle zutreffenden Zeichen mit linker Maustaste anzuklicken. Am Ende ihrer Bearbeitung sollte sie sofort den Button 'bestätigen!' aktivieren, der dann den zweiten Durchgang in analoger Form hervorruft. Grundsätzlich hinterlässt der Mausclick auf ein Zeichen keinerlei Feedback oder Markierung. Irrtümlich erneutes Anklicken auf ein korrektes Zeichen ergibt keinen Fehler, sondern kostet nur Testzeit. Nach der gesamten Testbearbeitung erhält die Testperson eine deskriptive Auswertung ihrer Ergebnisse.

Als Konzentrationsmaße dienen

- **d2c_genauigkeit:** Anzahl der insgesamt korrekt angeklickten zutreffenden Zeichen - Anzahl der Verwechslungen. [maximaler Wert=106]
- **d2c_zeit:** Die gesamte benötigte Testzeit in Sekunden
- **d2c_leistung:** $d2_genauigkeit/d2_zeit$

d2c_Genauigkeit misst die fehleradjustierte Sorgfalt bzw. Leistungsgüte oder Leistungsqualität. Die Anzahl der korrekt angeklickten Zeichen erfasst durch die Differenz zu den möglichen zutreffenden Zeichen zugleich die Auslassungen und jede Verwechslung ergibt einen Abzug. Da insgesamt wenige Fehler gemacht

wurden, erbringt der Fehlerabzug der falsch angeklickten Items jedoch kaum einen Vorteil. *d2c_leistung* setzt die fehleradjustierte Sorgfalt ins Verhältnis zur benötigten Zeit und vermittelt anschaulich die Anzahl korrekter Eingaben pro Sekunde.

Während in der Papierfassung des d2 als entscheidendes Konzentrationsmaß die Anzahl der korrekt markierten Zeichen bei konstanter Bearbeitungszeit anzusehen ist, gilt in der hier realisierten Computerfassung d2c als entscheidende Konzentrationsmessung das Arbeitstempo bei konstanter Anzahl von zu bearbeitenden Zeichen.

Bei der Auswertung der Daten stellte sich heraus, dass etliche Probanden (ca. 20-25%) unterschiedliche, und zum Teil massive Fehler produzierten, welche entweder ihren vollständigen Ausschluss aus diesen Tests erforderte bzw. nur die Wertung eines Testdurchgangs zuließ. In manchen Fällen war offenbar die Testanweisung falsch verstanden worden, in anderen Fällen wurde ein Durchgang einfach übersprungen oder eine hohe Fehleranzahl deutete auf einen klaren Ausreißer hin. Es ist schwer abzuschätzen, ob diese Fehler unzureichenden Bearbeitungshinweisen, softwareergonomischen Schwächen oder mangelnder Aufmerksamkeit und Motivation der Studierenden anzulasten sind.

Nachfolgende Ergebnisdarstellung in Tabelle 10 bezieht deshalb nur Studierende mit vollständig akzeptablen Testwerten in beiden Durchgängen einer Testung in die deskriptive Analyse ein. Die Berechnung der Retestreliabilitäten und der Mittelwertsunterschiede berücksichtigt nur Personen mit gültigen Angaben zu beiden Testzeitpunkten.

Tabelle 10: Ergebnisse zum d2c (Zeitpunkt 1: N=52; Zeitpunkt 2 N=53)

	α	M	s	Md	Schiefte	Kurtosis
<i>d2c_genauigkeit</i> :	.74	100.3	4.7	101.5	-1.10	.48
<i>d2c_genauigkeit</i> :	.48	101.3	3.8	102	-0.81	-.39
<i>d2c_zeit</i>	.96	148.5	21.3	144.7	.38	-.34
<i>d2c_zeit</i>	.90	135.8	15.6	135.7	.33	-.49
<i>d2c_leistung</i>	.93	.69	.10	.69	.31	-.17
<i>d2c_leistung</i>	.92	.75	.10	.69	.21	-.12

Retestreliabilitäten (Zeitabstand 8-12 Tage; N=40-41)

d2c_genauigkeit: .53
d2c_zeit : .80
d2c_leistung : .79

Mittelwertsunterschiede: $M_{Diff} = M_{\text{Zeitpunkt 2}} - M_{\text{Zeitpunkt 1}}$
t-test für abhängige Stichproben $df=38-39$

	M_{Diff}	t	pz	d
<i>d2c_genauigkeit</i>	0.88	1.3	.20	.20
<i>d2c_zeit</i>	-13.8	-6.9	.000	.74
<i>d2c_leistung</i>	0.73	7.2	.000	.75

Wie aus Tabelle 10 hervorgeht, erzielen *d2c_zeit* und *d2c_leistung* ansprechend hohe Reliabilitätskoeffizienten. Beide Maße korrelieren $r=-.93$ ($-.95$) miteinander und messen weitgehend dasselbe, weil sich in der *d2c_leistung* hauptsächlich der Einfluss der Bearbeitungszeit manifestiert und nur wenige Fehler vorkamen (im

Durchschnitt: 4.5% bzw. 3.3% Auslassungen sowie 0.6 bzw. 0.5% Verwechslungen. Zudem sind *d2c_genauigkeit* und *d2c_zeit* unabhängig voneinander ($r = -.02$, $r = -.16$). Der Mittelwert in *d2c_leistung* entspricht der durchschnittlichen Anzahl der fehleradjustierten, korrekt angeklickten Zeichen pro Sekunde. 1 geteilt durch *d2c_leistung* gibt die Zeit in Sekunden an, die benötigt wird, um Fehler adjustiert ein korrektes Zeichen zu markieren.

Bei der zweiten *d2c*-Testung, ca. 8 bis 12 Tage nach der ersten Messung erzielten die Studierenden signifikant höhere Konzentrationsleistungen hinsichtlich der *d2c_zeit* und *d2c_leistung*, die einer mittleren bis hohen Effektstärke entsprechen. Bezogen auf das Ausgangsniveau im Vortest *d2c*- hat sich die Bearbeitungszeit bzw. Leistung im Nachtest um ca. 8 bzw. 9 Prozent verbessert und fällt damit etwas geringer aus als der Wiederholungseffekt bei Bühner et. al. (2006) mit dem konventionellen *d2*. Je schwächer die Ausgangswerte zu Beginn (Vortest), desto höher fiel der Konzentrationszuwachs bei der Wiederholung (Nachtest) aus (*d2c_zeit*: $r_s V, (V-N) = .59$; $N = 40$, $p < .001$). Allerdings ließ sich diese Beziehung bei der *d2c_leistung* nicht signifikant bestätigen. ($d2c_leistung$ $r_s V, (N-V) = -.25$).

Bereits im Verlauf der ersten *d2c*-Testung ließen sich Übungseffekte nachweisen. Wie oben bereits erwähnt, setzt sich eine *d2c*-Testung aus insgesamt zwei wie in Abbildung 6 dargestellten Durchgängen (Hälften) zusammen. Bei der Messung zum Zeitpunkt 1 fielen sowohl die Genauigkeit wie die Bearbeitungszeit im zweiten Durchgang hochsignifikant günstiger aus als im ersten Durchgang. (Genauigkeit: $t(51) = 3.88$, $d = .53$; Zeit: $t(51) = 3.77$, $d = .19$; p zweiseitig jeweils $< .001$), wengleich der Effekt bei der entscheidenden Bearbeitungszeit im niedrigen Effektstärkebereich liegt.

Zusammenfassung der Trainingsstudien

Wie die vorhergehenden Ausführungen nahelegten, hängt die Größe eines Trainingseffekts auch vom speziellen Konzentrationstest ab (siehe dazu auch Westhoff & Dewald (1990)). Diese Behauptung soll an dieser Stelle nochmals näher analysiert, graphisch veranschaulicht, sowie auch statistisch belegt werden, wengleich die bisherigen Befunde hinreichend deutlich ausfielen. Denn wenn der Trainingseffekt beim GU einer Effektstärke von $d = 1.99$, beim ZRF_20 aber nur einem $d = 0.50$ entspricht und die Übung beim GU eine Zeitersparnis von 45%, beim ZRF_20 aber nur von 10% bewirkt, dann sind solche Unterschiede hinreichend offensichtlich.

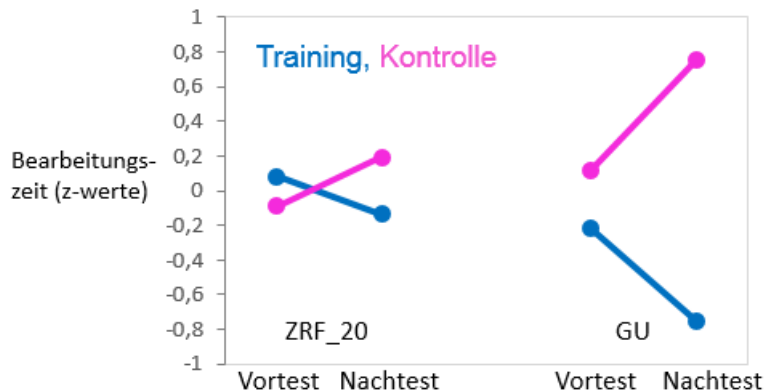
Um die deutlich unterschiedlichen Bearbeitungszeiten beider Tests vergleichbar zu machen, wurden die entsprechenden Zeiten zunächst für jeden Zeitpunkt gesondert z -transformiert. Dadurch werden u.a. auch die Mittelwertsunterschiede zwischen Vor- und Nachtest eliminiert. Mit den so modifizierten Daten wurde dann eine 3 faktorielle VA mit den Faktoren Zeitpunkt (Vortest, Nachtest), experimentelle Bedingung (Training, kein Training) und Konzentrationstest (ZRF_20, GU) berechnet. Hierbei dient die experimentelle Bedingung als between- und die restlichen Faktoren als within-Faktoren.

Die unterschiedliche Trainingswirkung beider Konzentrationstests müsste sich in der Wechselwirkung zweiter Ordnung zeigen und nur diese wurde daher geprüft. Dabei konnten nur die Probanden berücksichtigt werden, die vollständige akzeptable Daten für alle Faktoren aufwiesen. Es ergab sich ein deutlicher signifikanter Effekt für die Wechselwirkung zweiter Ordnung $F(1,53) = 38.04$, $p < .001$, der einem Eta-quadrat von 0.481 entspricht. Das gleiche Ergebnis kommt zustande, wenn man eine zweifaktorielle VA mit den Faktoren Konzentrationstest und den Differenzen aus Vor- und Nachtest berechnet und dann die einfache Wechselwirkung heranzieht.

Abbildung 7 zeigt die Mittelwerte aller verfügbaren Probanden für mindestens eine Trainingsstudie (= insgesamt 10 Probanden mehr als bei der statistischen Berechnung).

Abbildung 7

Interaktion zweiter Ordnung zwischen
Testzeitpunkt x experimenteller Bedingung x Konzentrationstest
 $F(1,53)=38,04$ $p<.001$



Die Grafik stellt keine klassischen Mittelwerte dar (siehe dazu Abb. 2 und 4), sondern lediglich einen fairen Vergleich für die unterschiedliche Trainingswirkung des ZRF_20 und des GU. Durch Training sinkt bei beiden Konzentrationstests die Bearbeitungszeit, aber das Auseinanderdriften von EG und KG fällt beim GU eben immens höher aus. Es gelingt kaum, durch Training die Reihenfolge ungeordneter Zahlen in aufsteigender Folge schneller anzugeben, aber man erzielt massive Übungsgewinne, aus einer Menge zweistelliger Zahlen solche zu identifizieren, deren erste Ziffer gerade und deren zweite Ziffer ungerade ist.

Tabelle 11 analysiert die Ergebnisse noch von einer etwas anderen, versuchsplan-technisch allerdings schwächeren Perspektive aus, in dem nur die Veränderungen von Vor- zum Nachttest unter der jeweiligen Bedingung betrachtet werden und hierbei auch der d2c Berücksichtigung findet.

Tabelle 11: Differenz (Nachttest - Vortest) für die Bearbeitungszeiten in Effektstärke für **Gruppe A** und **Gruppe B** unter KG (=kein Training) und EG-Bedingung (Training: 10 Übungstests) sowie unter **d2c als KG für alle Gruppen**

	KG	EG
GU:	.74	1.80
ZRF_20:	.10	0.47
d2c:	.66	

Anmerkung: Positive Effektstärken bedeuten Konzentrationszuwachs. Als relevante Streuung zur Berechnung der Effektstärke beim GU ($s=22.29$) und ZRF_20 ($s=4.5$) dienen hier die Kennwerte bei Jacobs 2013, da sie auf einem höheren N basieren und auf einer einmaligen Testung basieren. Für den d2c gilt die Vorteststreuung ($s=20.69$). Die gewählten Streuungen führen zu geringeren Effektstärken als die übliche Berechnung, wie sie weiter oben zum Einsatz kam.

Gruppe A erzielte unter KG-Bedingung im GU ($d=.74$) einen höheren Konzentrationszuwachs als unter EG-Bedingung im ZRF_20 ($d=.47$), was anfangs unter der Prämisse "vergleichbare Übungsgewinne der Tests" nicht erwartet wurde. Bei

Gruppe B hingegen stimmt die ursprüngliche Erwartung, im Training jeweils mehr Fortschritte als im Test ohne Übung zu erzielen, da sie unter Kontrollbedingung im ZRF keinen Gewinn ($d=.10$), in ihrer EG-Bedingung aber einen riesigen Vorteil im GU ($d=1.8$) erzielte. Für beide Gruppen diente der d2c als zusätzliche Kontrolle und erbrachte für Gruppe A und B zu beiden Zeitpunkten vergleichbare Testgewinne. Die Ergebnisse unter der jeweiligen KG in Tabelle 11 deuten darauf hin, der ZRF_20 sei auch übungsresistenter als der d2c einzuschätzen.

Transferiert die Übung eines Konzentrationstests auf die Leistung im anderen Konzentrationstest?

Während das mehrfache Üben stets die Leistung im eingeübten Konzentrationstest signifikant verbesserte, fiel die Verbesserung im GU unter Kontrollbedingung höher aus als der Übungsgewinn bei ZRF_20 unter EG-Bedingung. Es erscheint allerdings ziemlich abenteuerlich, für den Leistungsvorteil des GU unter der Kontrollbedingung das Training des ZRF_20 verantwortlich zu machen, weil der Leistungsvorteil der KG im Nachtest vollständig durch die Vortestung erklärbar ist. Zwar könnte die mehrfache Computertestung in der Übung die Gewöhnung an das Prozedere eines Testablaufs verstärken und das Handling mit der Maus verbessern, der so erzielte Transfereffekt auf nicht eingeübte Tests muss allerdings als verschwindend gering eingeschätzt werden. Eine letztlich befriedigende empirische Klärung dieser Frage hätte zwei weitere randomisierte Gruppen erfordert, die ausschließlich die Übungsserie eines der beiden Tests und erst im Nachtest alle 3 Konzentrationstests bearbeitet hätten.

Da aber für beide Tests gilt, dass die jeweilige Veränderung vom Vor- zum Nachtest unter Kontrollbedingung hoch vergleichbar der jeweiligen Veränderung vom Vortest zum ersten Übungstest unter EG-Bedingung ausfällt, bleibt als bestmögliche Erklärung für die Befunde nur die Schlussfolgerung, wiederholte Testung eines Konzentrationstests verbessere die Leistung in diesem Test, transferiere aber nicht auf die Konzentrationsleistung des anderen Tests. Somit wurde durch wiederholte Testung nicht die Konzentrationsfähigkeit, sondern lediglich eine spezielle Fertigkeit gefördert.

Mögliche Erklärungen der unterschiedlichen Übungswirkung von GU und ZRF_20

Vermutlich erfordert es zu Beginn einige Denkarbeit, die einzelnen Zahlen beim GU auf Ziel- oder Distraktoreigenschaft hin zu analysieren, weil die Teilbarkeit einer Ziffer durch 2 ziemlich abstrakt erscheint. Eine Zahl mag direkt wahrnehmbar sein, aber es ist zu Beginn recht ungewöhnlich, die Ziffern einer zweistelligen Zahl darauf hin zu prüfen, ob sie durch 2 teilbar sind und spontan eine Zahl als "erst gerade, dann ungerade" wahrzunehmen. Beim GU sind zu Beginn vermutlich eher Personen bevorteilt, die häufiger mit Zahlen und deren Eigenschaften umgehen. Mit zunehmender Übung gelingt es offenbar immer besser, eine wirksame Strategie zur Analyse der Teilbarkeit aufzubauen, Prüfprozesse zu vereinfachen und im weiteren Übungsverlauf manche Zielzahlen sowie offensichtliche Distraktoren sofort zu erkennen. Die relativ hohen Interkorrelationen der Bearbeitungszeiten im Verlauf der Trainingsübungen (siehe Anhang 3b) deuten auf vergleichbar effektive Strategien der Probanden hin. Anfänglich schwache Probanden haben bei der ersten Testung noch gewisse Probleme, holen dann aber zügig auf. Schließlich lassen sich die Zielitems immer einfacher global erfassen und der Weg bis dorthin ist mit stetigen Übungsgewinnen verbunden. Um den zu Beginn starken Übungseffekt beim GU etwas einzudämmen, wäre es eventuell günstiger, zunächst einen für das Testergebnis nicht gewerteten Übungsdurchgang voran zu stellen. Eine analoge

Maßnahme könnte auch die anfängliche Testwirkung des d2c etwas eindämmen, was in der Papierversion des d2 ja dadurch ansatzweise auch geschieht, als die erste Zeile nicht gewertet wird.

Die Zahlen 1 bis 20 sind für jeden Probanden unmittelbar wahrnehmbar. Da eine Zahlenfolge gedanklich sofort präsent ist, erfordert es auch keine Denkarbeit nach Anklicken auf die 4 nach der 5 zu suchen. Das Problem besteht vielmehr darin, die nächste Zahl auch zu finden. Da alle Zahlen stets nach Zufall angeordnet werden, kann man sich nicht an bestimmten Positionen oder Beziehungen der Zahlen untereinander orientieren, sondern muss stets den gesamten Zielraum absuchen. Hier böte sich an, bei der Suche nach der Zielzahl gleichzeitig die Positionen nachfolgender Zahlen mit zu verfolgen und sich diese möglichst einzuprägen. Vermutlich gehen viele Probanden auch so vor, aber die dafür notwendige Enkodierung kostet auch Zeit und legt diesem Vorgehen gewisse Grenzen auf, die weitere Übungsgewinne auf diesem Wege schnell limitieren. Es scheint offenbar sehr schwierig zu sein, durch reine Übung eine durchgreifend effizientere Strategie aufzubauen als diejenige, welche man ohnehin anwendet. Das Ergebnis hängt somit in höherem Maße von der Konzentration ab.

Während die Studie eindeutig den schnellen Leistungsfortschritt beim GU bestätigt hat, muss offen bleiben, ob sich der schwache Übungsgewinn beim ZRF_20 durch weitere massive Übungen nicht noch deutlicher steigern lässt. Beide Testverfahren sind von der reinen Gesamtübungszeit nicht direkt vergleichbar, da der GU bei der ersten Testung insgesamt ca. 3-mal so viel Bearbeitungszeit beansprucht wie der ZRF_20 (Berechnung auf das Basis der Daten von Jacobs 2013). Aber selbst bei vergleichbarer Gesamtübungszeit wäre der ZRF_20 höchstwahrscheinlich übungsresistenter als der GU.

Validitätshinweise für die Konzentrationstests

Fast alle Studierenden dieser Studie nahmen auch an einem späteren Experiment zur Wirkung von Lärm auf die Konzentrationsleistung teil und bearbeiteten dort einen weiteren Computerkonzentrationstest, den REV_c (siehe Anhang 4 bzw. Jacobs 2014). Dieser basiert auf einer Idee des Revisionstests REV_T von Marschner (1972) und verlangt von den Testpersonen, aus einer Menge einfacher einstelliger Additionsaufgaben mit richtigem oder falschem Ergebnis, diejenigen mit korrekter Summe anzuklicken. Der REV_c wurde einmal mit und einmal ohne Lärm zur Bearbeitung vorgelegt und erzielte zu zwei Testzeitpunkten Konsistenzen über .90. Da der Lärm keinen Einfluss auf die Konzentrationsleistung ausübte, geht der Durchschnitt aus beiden Testungen als sehr zuverlässiges Maß für den REV_c in die weitere Analyse ein.

In der Studie von Jacobs (2013) korrelierten die Bearbeitungszeiten von GU - und ZRF_20 untereinander mit $r = .50$. Beide Tests standen in ähnlich hohem Zusammenhang mit einer Onlinevariante des Konzentrationsleistungstest frei nach Düker & Lienert (2001). Als weitere Validitätshinweise wurden entsprechend hohe Zusammenhänge mit dem d2c und dem REV_c erwartet. Als relevante Konzentrationsmaße gelten nur die Bearbeitungszeiten bzw. diejenigen Konzentrationsleistungen, welche im Wesentlichen das Arbeitstempo erfassen. Wie aus Tabelle 12 hervorgeht, korrelieren alle Tests signifikant untereinander in erwarteter Richtung.

Tabelle 12: Interkorrelationen der Konzentrationsmaße. (N: 49-57)

	ZRF_20	d2c	REV_c
GU	.44	.50	-.64
ZRF_20		.62	-.51
d2c			-.57

Anmerkung: Alle Korrelationen sind auf dem 0.1% Niveau signifikant.

Wählt man für GU und d2c statt der in Tabelle 12 verwendeten Bearbeitungszeit die Konzentrationsleistungen (Genauigkeit/Zeit), so erhält man hoch vergleichbare Ergebnisse. Neben einer Bestätigung des bereits von Jacobs nachgewiesenen signifikanten Zusammenhangs zwischen GU und ZRF_20 und der erneuten signifikanten Bestätigung hier, erhalten diese beiden Tests weitere Validitätsbelege durch die relativ hohen Zusammenhänge mit dem d2c sowie dem REV_c von jeweils mindestens $r=.50$. Eine Faktorenanalyse der 4 Tests ergab einen Faktor, der insgesamt 69% der Varianz erklärt und auf dem alle Tests Ladungen zwischen .79 bis .87 aufweisen.

Da es theoretisch plausibel erscheint anzunehmen, gute Schulleistungen seien eher bei hoher Konzentrationsfähigkeit zu erzielen, waren entsprechende Korrelationen zwischen allen Konzentrationstests und dem Abiturnotendurchschnitt erwartet worden. Entgegen der Erwartung ließen sich hier aber keinerlei signifikante Zusammenhänge zwischen allen möglichen Messvariablen der 4 Konzentrationstests und dem Abiturnotendurchschnitt feststellen. Somit ist es nicht gelungen, die von Jacobs 2013/2013b berichteten signifikanten Zusammenhänge zwischen einzelnen Konzentrationsmaßen des ZRF_20 sowie des GU und dem Abiturnotendurchschnitt zu replizieren. Des Weiteren konnten bei keinem der Konzentrationstests unterschiedliche Leistungen für die Geschlechter nachgewiesen werden, womit auch die von Jacobs 2013 berichteten signifikanten Vorteile für Männer beim GU keine Bestätigung erfahren.

Messen die Tests tatsächlich Konzentration?

Alle Tests erfüllen die für Konzentrationstests gültigen Regeln nach Westhoff und Hagemeister (2005)

- Die Wahrnehmung der Reize ist ganz eindeutig.
- Die Aufgaben erfordern keine anspruchsvollen Gedächtnisprozesse und basieren auf einer einfach zu erinnernden Regel.
- Die Items stellen keine komplexen Probleme da und verlangen keine hohe Lernfähigkeit.

Alle Tests, die nach einer Studie von Schmidt-Atzert, Bühner und Enders (2006) hoch auf dem Faktor Konzentration laden, weisen nach Ansicht der Autoren auch formale Gemeinsamkeiten auf, die ebenfalls auf die hier behandelten Tests zutreffen: "Sie geben einfache und gut wahrnehmbare Reize vor, verlangen die Anwendung von einfach zu merkenden Regeln sowie die Koordination von Teileroperationen (z. B. Identifizieren, Vergleichen, Durchstreichen) und die Testleistung wird über das Arbeitstempo operationalisiert." Während bei Papier und Bleistift-Tests alle Testbögen eines Tests identisch ausfallen, ist dies bei den eingesetzten Onlinetests nicht vollständig gewährleistet, da die Darstellung der Reize unter anderem vom Browsertyp, der Bildschirmgröße und dem eingestellten Zoom abhängig ist. Geht man jedoch von der plausible Annahme davon aus, der

Studierende habe seinen Browser so eingestellt, wie er sich normalerweise Seiten im Internet ansieht, so liefern die Onlinetests auf jeden Fall eine brauchbare Grundlage, die, wenn auch nicht optimale, so doch eine hinreichende Wahrnehmung ermöglicht. Es macht offensichtlich auch wenig aus, dass die Onlinetest jeweils Zufallsanordnungen aus der Menge möglicher Zeichen repräsentieren, da diese Auswahl offensichtlich die Kontentvalidität gewährleistet sowie durch die Reliabilität belegt wird.

Nach Ansicht der Autoren müssen Konzentrationstests sich zudem von anderen Faktoren abgrenzen lassen. Sie sollten eben Konzentration und nicht Intelligenz, Gedächtnis, Reaktionsschnelligkeit oder motorische Bearbeitungsgeschwindigkeit messen. Im Rahmen faktorenanalytischer Studien konnten Schmidt-Atzert et al. (2006) einen Konzentrationsfaktor extrahieren, der relativ unabhängig von etlichen unerwünschten Einflüssen war, obgleich das Reizmaterials eines Konzentrationstests (numerisch oder figural) einen gewissen Einfluss ausübte.

Alle hier zum Einsatz gekommenen Onlinetests waren mit der Maus zu beantworten und setzen eine ausreichende Geschicklichkeit bei der Mausnavigation und dem Anklicken der linken Maustaste voraus. Wenngleich Daten fehlen, um zweifelsfrei abschätzen zu können, wie stark die Geschicklichkeit im Umgang mit der Maus in das Konzentrationsergebnis eingeht und so die Validitätskoeffizienten womöglich künstlich erhöht, bleibt zu hoffen, alle Studierenden verfügten über hinreichende Erfahrung der Mausnutzung und diesbezügliche Personenunterschiede fielen nicht besonders ins Gewicht.

Schmidt-Atzert et al. (2006) konnten durch Faktorenanalysen mit etlichen Konzentrationstests weiterhin aufzeigen, dass der REV-T, der d2 sowie der Zahlen-Verbindungs-Test (ZVT Oswald & Roth (1987)) hoch auf einem Konzentrationsfaktor und nur mäßig auf einem Materialfaktor laden sowie kaum etwas mit Intelligenz zu tun haben. Wegen der hier gefundenen Validitätskoeffizienten sowie der sehr hohen Korrelation des ZRF_20 mit einer Computerversion des Zahlen-Verbindungs-Test (ZVT) von $r_s = .81$ (siehe Anhang 5 bzw. Jacobs 2013b) kann überhaupt kein Zweifel daran bestehen, dass auch der ZRF_20 in hohem Maße Konzentration erfasst, die Computerversion des ZVT diese aber noch deutlich reliabler misst. Gegenüber dem GU, der durch die ansprechenden Korrelationen mit REV_c und d2c ebenfalls hauptsächlich Konzentration messen dürfte, hat der ZRF_20 neben der kürzeren Bearbeitungszeit den großen Vorteil, weniger anfällig gegenüber Trainingseffekten zu sein.

Literatur:

Brickenkamp, R. Schmidt-Atzert, L., Liepmann, D. (2010). d2 –Revision Aufmerksamkeits- und Konzentrationstest - Manual. Hogrefe

Bühner, M. Ziegler, M. , Bohnes, B. & Lauterbach, K. (2006). Übungseffekte in den TAP Untertests Test Go/Nogo und Geteilte Aufmerksamkeit sowie dem Aufmerksamkeits-Belastungstest (d2). *Zeitschrift für Neuropsychologie*, 17 (3), , 191–199.

Düker, H., Lienert, G. A. (2001) Konzentrations-Leistungs-Test Revidierte Fassung (KLT-R) Neubearbeitung von: Lukesch, H. & Mayrhofer, s. Hogrefe Verlag, Göttingen.

Hagemeister, C. (2005). Erkennen von "Übung in Konzentrationstests".
http://tu-dresden.de/die_tu_dresden/fakultaeten/fakultaet_mathematik_und_naturwissenschaften/fachrichtung_psychologie/i2/diagnostische/mitarbeiter/publikationen2/publikationen_carmen_hagemeister/ch_fgdi_2005.pdf

Jacobs, B. (2002). 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

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. (2013b). Ähnlichkeit zwischen Zahlen-Verbindungs-Test (ZVT) und Zahlenreihenfolgetest_20 (ZRF_20).
http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/zrf_20_vs_zvt.html

Jacobs, B. (2014). Beeinträchtigt Lärm die Leistung in einem Konzentrationstest?
URN: urn:nbn:de:bsz:291-psydok-49581
URL: <http://psydok.sulb.uni-saarland.de/volltexte/2014/4958/>

Marschner, G. (1972) Revisions-Test (Rev.T.). Göttingen: Hogrefe.

Oswald W, Roth E (1987). Der Zahlen-Verbindungs-Test (ZVT) –Handanweisung. Göttingen, Hogrefe.

Schmidt-Atzert, L. Bühner, M. & Enders, P. (2006). Messen Konzentrationstests Konzentration? Eine Analyse der Komponenten von Konzentrationsleistungen. Diagnostica, 52(1), 33-44.

Westhoff, K. (1989). Übungsabhängigkeit von Leistungen in Konzentrationstests. Diagnostica, 35 (2) 122-130.

Westhoff, K. & Dewald, D. (1990). Effekte der Übung in der Bearbeitung von Konzentrationstests. Diagnostica 36 (1) 1-15.

Westhoff, K. & Hagemeyer, C. (2005). Konzentrationsdiagnostik. Pabst, Lengerich.

Anhang

Anhang 1: Aufgabenbeispiel eines Durchgangs des GU-Konzentrationstests

**Erste Ziffer gerade
zweite Ziffer ungerade!**

18	39	85	28	49	69	68	23	31	16
26	35	43	34	15	71	23	31	63	68
46	58	63	47	38	87	66	74	67	68
39	84	64	41	43	96	31	62	64	82
57	68	38	28	38	21	79	68	52	24
24	83	42	16	18	27	47	43	53	86
17	28	18	68	45	16	67	79	64	53
98	38	13	45	98	69	91	26	48	61
95	68	86	24	75	95	68	75	83	32
87	44	26	66	96	16	93	28	78	36

Bearbeitung bestätigen!

Anhang 2. Beispiel eines Durchgangs des ZRF_20-Konzentrationstest

Klicke die Zahlen beginnend von 1 in der richtigen Reihenfolge an!

19	13	11	5	6
4	12	15	7	16
8	2	9	3	17
18	14	10	1	20

Anhang 3 a)

GU: Interkorrelationen der Bearbeitungszeiten über alle Messzeitpunkte (N=22-30)												
	Vortest	1	2	3	4	5	6	7	8	9	10	Nachtest
Vortest		.61	.58	.55	.67	.61	.56	.63	.50	.57	.54	.54
1	.61		.92	.74	.76	.57	.56	.58	.50	.73	.77	.67
2	.58	.92		.77	.79	.67	.64	.67	.45	.75	.79	.72
3	.55	.74	.77		.91	.81	.82	.79	.56	.75	.78	.79
4	.67	.76	.79	.91		.87	.79	.83	.57	.84	.82	.83
5	.61	.57	.67	.81	.87		.82	.90	.56	.85	.86	.79
6	.56	.56	.64	.82	.79	.82		.89	.63	.73	.83	.87
7	.63	.58	.67	.79	.83	.90	.89		.56	.87	.92	.93
8	.50	.50	.45	.56	.57	.56	.63	.56		.85	.82	.61
9	.57	.73	.75	.75	.84	.85	.73	.87	.85		.89	.86
10	.54	.77	.79	.78	.82	.86	.83	.92	.82	.89		.90
Nachtest	.54	.67	.72	.79	.83	.79	.87	.93	.61	.86	.90	

Anhang 3b)

ZRF_20: Interkorrelationen der Bearbeitungszeiten: alle Messzeitpunkte(N=24-31)												
	Vortest	1	2	3	4	5	6	7	8	9	10	Nachtest
Vortest		.76	.66	.58	.69	.59	.82	.72	.68	.68	.50	.67
1	.76		.86	.41	.65	.66	.88	.77	.66	.62	.70	.59
2	.66	.86		.46	.67	.59	.82	.79	.69	.69	.75	.64
3	.58	.41	.46		.59	.40	.56	.56	.47	.83	.59	.47
4	.69	.65	.67	.59		.70	.82	.85	.90	.69	.76	.73
5	.59	.66	.59	.40	.70		.76	.64	.74	.60	.63	.57
6	.82	.88	.82	.56	.82	.76		.85	.85	.80	.84	.73
7	.72	.77	.79	.56	.85	.64	.85		.75	.75	.70	.84
8	.68	.66	.69	.47	.90	.74	.85	.75		.59	.81	.65
9	.68	.62	.69	.83	.69	.60	.80	.75	.59		.77	.67
10	.50	.70	.75	.59	.76	.63	.84	.70	.81	.77		.65
Nachtest	.67	.59	.64	.47	.73	.57	.73	.84	.65	.67	.65	

Anhang 4: Beispiel eines Durchgangs des REV_c

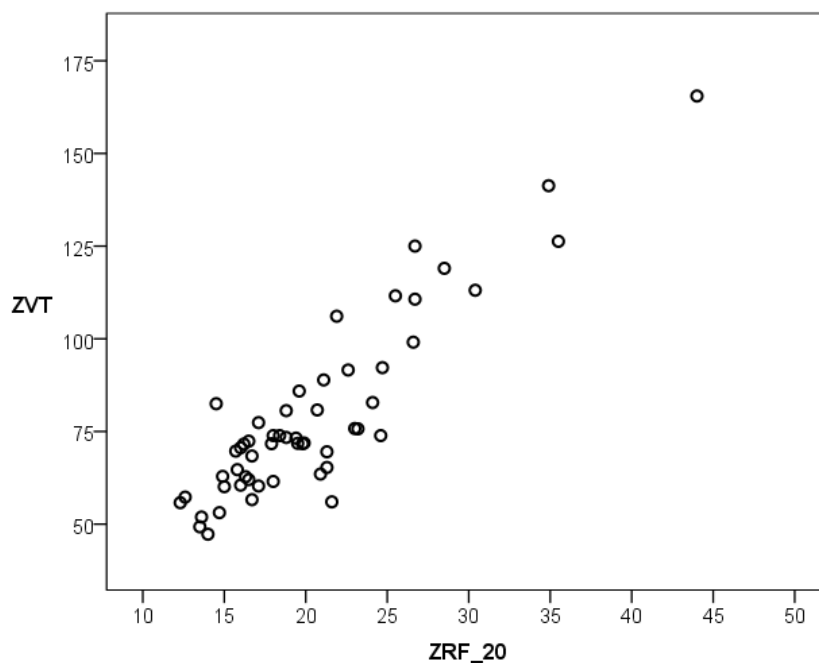
Klicke alle Aufgaben mit korrektem Ergebnis an!

Durchgang 1

1+2=4	1+5=7	2+5=8	1+3=5	2+3=6	2+1=4
4+3=8	3+1=5	5+1=6	1+4=6	4+4=8	3+5=8
1+4=6	4+4=9	1+4=6	4+5=10	3+5=8	1+5=6
4+5=10	5+2=8	2+5=8	2+1=4	4+3=8	1+4=6
2+3=6	3+4=7	5+1=7	2+1=4	4+1=6	1+3=5
5+4=10	1+3=5	5+2=8	4+5=10	5+3=9	5+3=8
3+4=8	1+2=3	3+1=4	5+4=10	1+1=3	2+4=6
2+5=7	3+3=7	1+3=5	2+2=5	4+4=8	5+5=11
3+1=4	2+1=3	2+2=5	2+4=6	1+4=6	1+3=4
1+5=6	5+4=10	5+4=10	4+5=10	4+4=8	3+3=7
3+4=7	2+4=7	1+4=5	5+2=7	4+4=8	4+2=7
5+1=6	3+4=8	1+1=2	3+4=8	2+2=5	3+2=6

bestätigen!

Anhang 5: Zusammenhang zwischen Zahlen-Verbindungs-Test und ZRF_20 - durchschnittliche Bearbeitungszeiten pro Durchgang in Sekunden -



Die Korrelation zwischen ZRF_20_zeit und ZVT-zeit beträgt $r = .91$ ($r_s = .81$)
[N=54]. Aus:

http://bildungswissenschaften.uni-saarland.de/personal/jacobs/diagnostik/tests/konzentration/zrf_20_vs_zvt.html