

inspective.

**Replikation einer Studie zur Verknüpfung von Wahrnehmung und
kognitiver Leistungsfähigkeit**

Marc Dressler

Stuttgart,
Februar 2015

Zusammenfassung

Wahrnehmung und Denken sind in der Kognitionspsychologie als Informationsverarbeitung konzeptualisiert. Beim Denken lassen sich logische und kreative Leistungen unterscheiden; bei der Wahrnehmung, je nach Detailtiefe, eine globale Verarbeitung, die figürliche Zusammenhänge ins Bewusstsein hebt, und eine lokale Verarbeitung, bei der elementare Details bewusst werden. Zur Kopplung von Wahrnehmung und Denken sind verschiedene Modelle vorgeschlagen worden, u.a. GLOMO^{sys} von Jens Förster, das postuliert, dass sich das logische Denken bei einer lokalen Verarbeitung verbessert und sich bei einer globalen Verarbeitung verschlechtert, wohingegen sich das kreative Denken bei einer lokalen Verarbeitung verschlechtert und bei einer globalen Verarbeitung verbessert. Eine Studie Försters, in der das Modell mittels globalem und lokalem Priming gegenüber einer Kontrollgruppe für fünf Sinnesmodalitäten bestätigt wurde, geriet aber in den Verdacht der Datenmanipulation. Um zur Aufklärung des Verdachts beizutragen, sollte Försters umstrittene Studie in einem online-Experiment repliziert werden. Aus technischen Gründen beschränkte sich die Replikation auf die Kopplung der visuellen Wahrnehmung mit der kognitiven Leistung. An dem Experiment nahmen 37 Studierende teil im Alter von 18 bis 25 Jahren und ohne ersten Studienabschluss. Die Studierenden wurden zu Beginn des Experiments zufällig entweder der globalen, der lokalen oder der Kontrollgruppe zugeteilt und sowohl ihre kreative als auch ihre logische Leistung nach dem Priming psychometrisch erfasst. Das Experiment ergab keine signifikanten Unterschiede in den kognitiven Leistungen zwischen den Gruppen. Die gescheiterte Replikation ist zusammen mit aufgedeckten Unstimmigkeiten nicht dazu angetan, Förster vom Fälschungsverdacht zu entlasten.

Replikation einer Studie zur Verknüpfung von Wahrnehmung und kognitiver Leistungsfähigkeit

Es ist immer wieder verblüffend zu sehen, welche Ähnlichkeiten vorüberziehende Wolken haben können mit Menschen, Tieren, Fahrzeugen oder gar kompletten Szenerien. Die einzelnen Wolken verschwinden dann hinter dem zusammenhängenden Eindruck eines uns bekannten Musters. Der globale Aspekt der figürlichen Wahrnehmung dominiert über die lokalen Elementen, aus denen sich die Figur zusammensetzt.

Die natürliche Tendenz des Menschen zur Wahrnehmung von Mustern, deren Bildungsgesetze in der Gestaltpsychologie untersucht und beschrieben worden sind, findet ihre evolutionstheoretische Begründung darin, dass es überlebensrelevant sein kann, wiederkehrende Gefahrensituationen schnell zu erkennen und das dazu passende Verhaltensschema des Flüchtens oder Kämpfens abzurufen. Aber auch in weniger bedrohlichen Situationen ist ein globaler Überblick über eine Situation ökonomischer als ihre jedesmalige Rekonstruktion Stück für Stück aus den einzelnen Elementen.

Die Kehrseite dieser Wahrnehmungsökonomie ist, dass wir auch dort Muster wahrnehmen, wo gar keine sind. Diese Cluster-Illusion (Gilovich, Vallone & Tversky, 1985) kann durchaus pathologische Züge annehmen. Der schwedische Dramatiker August Strindberg notierte in seinem *Okkulten Tagebuch*, dass er im Kissen seiner Ruhestätte einmal einen Marmorkopf im Stile Michelangelos erkannte, ein andermal schreckliche Monster oder gotische Wasserspeier und eines Nachts begrüßte ihn gar der Satan persönlich. Der an Apophanie leidende Strindberg maß seinen illusorischen Wahrnehmungen der Kissenfalten schicksalhafte Bedeutung zu, er stellte sie in einen globalen Kontext.

Die Unterscheidung in global und lokal lässt sich demnach nicht nur treffen hinsichtlich der Wahrnehmung, sondern auch hinsichtlich des Denkens. Veranschaulichen lässt sich das für das Denken an der hermeneutischen Problemstellung, ob wir die Bedeutung eines Satzes begreifen aus der Bedeutung der einzelnen Wörter, oder ob wir umgekehrt die Bedeutung der einzelnen Wörter begreifen aus dem Verständnis des gesamten Satzes. Ist es der globale Sinn des Textes, der die Bedeutung der Sätze verständlich macht, oder sind es die lokalen Bedeutungen der Sätze, die uns den gesamten Text verstehen lassen?

Einem psychologischen Ansatz zur Erforschung des Verarbeitungsstils des Denkens (Förster, 2009) zufolge äußert sich globales Denken im kreativen Schließen, in dessen Folge Begriffe intensional erweitert werden, während lokales Denken im logischen Schließen zu

finden ist, in dessen Folge Begriffe intensional verengt werden. Beim logischen Schließen darf der Schlusssatz begrifflich nicht mehr enthalten als die Prämissen, aus denen er folgt; kreatives Denken ist dagegen gerade dadurch gekennzeichnet, dass sein Ergebnis reicher ist als sein begrifflicher Ausgangspunkt.

Diese Konzeptualisierung der Denkstile legte Förster einer Reihe von Versuchen zugrunde, die ihm zur Überprüfung seines Modells psychischer Informationsverarbeitung mit der Bezeichnung GLOMO^{sys} dient (siehe Abbildung 1). Im Kern besteht dieses Modell aus zwei getrennten Verarbeitungssystemen jeglicher Information, perzeptiver wie kognitiver: dem globalen System und dem lokalen System (Förster & Dannenberg, 2010). Aktiv ist entweder das eine oder das andere System. Das Betätigungsfeld des globalen Systems ist in GLOMO^{sys} die Neuheit; es verarbeitet mehrdeutige, unklare, komplexe oder abstrakte Informationen, um das Neue in den übergreifenden Erfahrungs- und Wissenshorizont einer Person zu integrieren. Das lokale System kommt dagegen bei Vertrautem zum Einsatz; es verarbeitet eindeutige, klare, einfache oder konkrete Informationen, um das Vertraute genauer zu differenzieren.

Laut GLOMO^{sys} sind der Verarbeitungsstil der Wahrnehmung und der des Denkens aneinander gekoppelt. Globale Wahrnehmung führt demnach zu globalem Denken, lokales Wahrnehmen zu lokalem Denken. Den Verarbeitungsstil auslösen können systemexterne Faktoren wie Farbe, Stimmung, Machtfülle, Religionszugehörigkeit oder eben der Grad des Vertrautseins bzw. der Neuheit. Entscheidend ist, dass der einmal aktivierte Wahrnehmungsstil sich überträgt auf andere psychische Vorgänge wie das kreative oder logische Denken. Diesen Umstand macht sich Förster zunutze, um mittels prozeduralem Priming bei Versuchspersonen gezielt über die Wahrnehmung die Aktivierung eines Verarbeitungssystems zu induzieren und anschließend deren kognitive Leistung zu messen.

GLOMO^{sys} zufolge müssten global geprimte Versuchspersonen bei kreativen Aufgaben besser abschneiden als lokal geprimte. Bei logischen Aufgaben ist es genau umgekehrt. In der Gruppe des lokalen Primings haben die Versuchspersonen die Aufgabe, die Elemente im Primingreiz zu identifizieren, während sie in der Gruppe des globalen Primings die Figur benennen müssen, die sich aus den Elementen zusammensetzt. Anhand dieses prozeduralen Primings kontrolliert Förster den Wahrnehmungs- und Denkstil seiner Versuchspersonen und kann im Anschluss die Effekte des Verarbeitungsstils messen.

Die Effekte, dass globales Priming im Mittel zu kreativeren Ergebnissen führt, dafür aber das logische Denken leidet, und dass lokal Geprimte beim logischen Denken punkten, dafür

aber bei kreativen Aufgaben Abstriche machen müssen, während Mitglieder der Kontrollgruppe sich jeweils im Mittelfeld behaupten, konnte Förster für die fünf Sinnesmodalitäten Sehen, Hören, Tasten, Schmecken und Riechen (Förster & Denzler, 2012) nachweisen. Diese Studie geriet allerdings, wie zwei weitere Studien (Förster, 2009 und Förster, 2011), in den Verdacht der Datenmanipulation.

Auffällig an ihnen ist, dass die Mittelwerte für die kognitiven Leistungen der Priming- und Kontrollgruppen annähernd auf einer Geraden liegen, während der Verlauf von Mittelwerten in Vergleichsarbeiten, die ebenfalls die genannten Primingeffekte dokumentieren, einen deutlichen Knick ausweist (Anonymous, 2012). Dass sich konsistent ein linearer Zusammenhang zwischen den Mittelwerten ergibt, ist statistisch extrem unwahrscheinlich. Eine Durchsicht der Originaldaten wäre daher angezeigt, doch die sind nicht mehr vorhanden.

An den veröffentlichten Daten fällt darüber hinaus der Frauenanteil an den studentischen Versuchspersonen auf. Der lag laut Försters Studien bei knapp über der Hälfte; an der Fakultät für Psychologie der Universität Amsterdam, an der Förster zu jener Zeit beschäftigt war, waren jedoch fast Dreiviertel Studentinnen eingeschrieben. So kam der Niederländische Rat für Forschungsintegrität zu dem Schluss, dass die Daten in Förster & Denzler (2012) manipuliert worden seien, woraufhin die Universität empfahl, die Studie zurückzuziehen.

Das lehnte Förster ab mit dem Hinweis, dass er die Untersuchungen zwischen 1999 und 2008 in Deutschland vorgenommen habe. eMails aus dem Jahr 2009, in denen es um die Vorbereitung einer dieser Untersuchungen geht (Kolschooten 2014), stehen laut Förster im Zusammenhang mit Plänen, einen Teil der Untersuchungen zu wiederholen. Für eine solche Wiederholung gibt es inzwischen mehr Gründe, als Förster lieb sein dürfte. Zumal eine weitere Studie von ihm (Friedman & Förster, 2002) nicht repliziert werden konnte (Wronska & Blazejewicz, 2013).

Um einen kleinen Beitrag zur Aufklärung des Falles zu leisten, sollte mit der hier vorge-tragenen Untersuchung der Versuch einer Replikation der Effekte in Förster & Denzler (2012) unternommen werden, die mittels visuellem Priming erzielt wurden. Das visuelle Priming wurde ausgewählt, weil die Untersuchung online ausgeführt werden musste und auf diesem Weg außer dem akustischen nur ein visueller Priming möglich ist, letzteres sich aber leichter kontrollieren lässt. Schließlich muss ein Teilnehmer, um bis ans Ende der Untersuchung zu gelangen, den Monitor eingeschaltet lassen, den Lautsprecher dagegen nicht.

Methode

Teilnehmerkreis

Die Untersuchung strebte eine Teilnahme von 60 bis 66 Versuchspersonen an. Die Anzahl ergibt sich bei Anwendung des Statistik-Programms G*Power3 (Faul et al., 2007), um bei Prüfung einer gerichteten Unterschiedshypothese bei einer Teststärke von .8 und $p = .05$ einen großen Effekt nachzuweisen. Damit die Stichprobe der Replikation vergleichbar ist mit der von Förster & Denzler (2012) (Asendorpf et al., 2013), wurde der Teilnehmerkreis eingeschränkt auf Studierende ohne ersten Studienabschluss im Alter von 18 bis 25 Jahren.

Rekrutiert wurden die Teilnehmer über direkte Ansprache per eMail, Einstellung in Internet-Portalen und -Foren wie Facebook oder Kleiderkreisel sowie Aushänge mit QR-Code, der zum Link der Untersuchung leitete, in Universitätsgebäuden. Der Link zur Untersuchung war freigeschaltet vom 15. bis zum 28. Dezember 2014.

Visuelles Priming

Das Priming der Verarbeitungsstile erfolgte über Navon-Letters (Navon, 1977), welche Buchstaben darstellen, die wiederum aus Buchstaben zusammengesetzt sind. Es wurden gemäß der Angaben in Förster & Denzler (2012) acht großen Buchstaben aus 5 mal 5 Buchstaben (L, H, E und F) der Größe .5 mal .5 cm konstruiert. Da Angaben zur Schriftart fehlten, wurde in Anlehnung an ein Buchstabenbeispiel aus Förster und Dannenberg (2010) die Standardschrift Times New Roman verwendet. Die angefertigten Buchstaben deckten sich sehr gut mit denen, die Förster per eMail nach ihrer Anfertigung zur Verfügung stellte.

Mit dem Seitenaufruf der Untersuchung wurde eine Versuchsperson zufällig der globalen, der lokalen oder der Kontrollgruppe zugeteilt.

Die Versuchspersonen der globalen Gruppe wurden anhand eines Beispielbuchstabens instruiert, per Klick zu entscheiden, ob es sich bei dem großen Buchstaben um ein L oder ein H handelt. Die vier Buchstaben (L und H jeweils bestehend aus Es und Fs) wurden aus technischen Gründen in einer festen Reihenfolge zwölf Mal präsentiert, nicht in zufälliger Abfolge wie bei Förster & Denzler (2012). Ebenfalls abweichend von der Originalstudie wurden die Buchstaben in Schwarz vor einem weißen Hintergrund dargeboten und nicht in Weiß vor einem schwarzen Hintergrund; statt per Tastendruck gelangten die Versuchspersonen durch Anklicken eines ‚Weiter‘-Buttons zum nächsten Buchstaben.

Die Versuchspersonen der lokalen Gruppe wurden analog aufgefordert, zu entscheiden, ob es sich bei den kleinen Buchstaben um ein L oder ein H handelt. Die vier Buchstaben (E und F jeweils bestehend aus Ls und Hs) wurden den Personen ebenfalls in einer festen Reihenfolge zwölf Mal angezeigt.

Die Versuchspersonen der Kontrollgruppe wurden ebenfalls anhand eines Beispiels angeleitet, zu entscheiden, ob es sich bei den eingeblendeten Buchstaben um ein L oder ein H handelt, wobei die Ls und Hs entweder groß oder klein sein konnten. Die acht Buchstaben (der globalen plus der lokalen Gruppe) wurden in einer festen Reihenfolge sechs Mal präsentiert. Zwischen der Darbietung der Buchstaben wurde in allen drei Gruppen für 500 ms ein Pluszeichen an der Stelle eingeblendet, an der die 48 Buchstaben erschienen.

Um sicherzustellen, dass das Priming erfolgreich war, d.h. dass die Versuchspersonen der Priminggruppen ihre Aufmerksamkeit auf die Wahrnehmung der großen bzw. der kleinen Buchstaben fokussiert hatten und nicht einfach zufällig ein L oder ein H anklickten, sollten 75 Prozent der Klicks beim Priming überzufällig richtig sein. Daher wurden Versuchspersonen, die mehr als fünf Buchstaben falsch anklickten, fallweise ausgeschlossen.

Logisches Denken

Die Leistung des logischen Denkens ermittelten Förster & Denzler (2012) anhand von vier Aufgaben der Form ‚Wenn $A < B$ und $C > A$, dann folgt für B und C?‘, für deren Bearbeitung die Versuchspersonen insgesamt vier Minuten Zeit hatten. Diese Aufgaben waren nicht enthalten in dem Material, das Förster per eMail zur Verfügung stellte. Verwendet wurden daher die Übersetzungen von vier logischen Problemen („Analytical Reasoning Items“) des Graduate Record Examination (GRE), auf den Friedman & Förster (2000) verweisen.

Kreatives Denken

Für die Ermittlung der Leistung des kreativen Denkens wurde der Cartoon eingesetzt, den Förster per eMail zur Verfügung gestellt hat. Er zeigt einen Hund auf einem Sofa. Zu diesem Cartoon sollten die Versuchspersonen einen möglichst originellen Titel ersinnen. Dafür hatten sie zwei Minuten Zeit. Vier unabhängige Künstler bewerteten die Originalität der eingetippten Titel auf einer Likert-Skala von 1 bis 7.

Versuchspersonen, die weder auf einen Titel kamen noch eine logische Aufgabe lösten, wurden fallweise ausgeschlossen, weil bei ihnen davon auszugehen ist, dass sie weniger unfähig als vielmehr lustlos, also nicht ernsthaft bei der Sache waren.

Stimmung

Vor dem Priming und nach den Aufgaben wurde wie bei Förster & Denzler (2012) die Stimmung der Versuchspersonen anhand von zehn Items mittels einer Kurzversion des Positive and Negative Affect Schedule (PANAS) gemäß Watson, Clark & Tellegen (1988) in der Übersetzung von Krohne et al. (1996) abgefragt. Damit sollte kontrolliert werden, ob die Stimmung das Priming des Verarbeitungsstils außer Kraft setzt (vgl. Gaspar & Clore 2002).

Die Versuchsteilnehmer konnten die Untersuchung nur fortsetzen, wenn sie auf allen zehn Items einen Wert markiert hatten. Eine Imputation fehlender Werte war daher nicht erforderlich.

Den Versuchspersonen wurde in der Begrüßung mitgeteilt, dass sie an der Replikation einer unter Fälschungsverdacht stehenden Studie mitwirken, und dass die einzelnen Aufgabenteile unabhängig voneinander seien. Diese Irreführung wurde am Ende der Untersuchung aufgelöst mit dem Hinweis, dass der differenzielle Einfluss des Primings auf die kognitive Leistung untersucht werden sollte und insofern doch ein Zusammenhang zwischen den Aufgabenteilen besteht.

Ergebnisse

Von den 136 erzeugten Datensätzen gingen 39 auf technische Artefakte zurück wie beispielsweise den Seitenaufruf durch eine Suchmaschine. Darüber hinaus wurden 34 Teilnehmer ausgeschlossen, die entweder das Priming abbrachen oder zu viele falsche Angaben im Verlauf des Primings machten. Weitere 20 Teilnehmer konnten nicht berücksichtigt werden, weil sie zu alt oder zu jung waren. Schließlich mussten aufgrund des Bildungsniveaus noch 6 Teilnehmer ausgeschlossen werden, so dass am Ende $N = 37$ Versuchspersonen, davon 26 weibliche (70%) mit einem Durchschnittsalter von 22.4 Jahren ($SD = 1.8$) in die Auswertung eingingen.

Um zu bestimmen, ob sich die Leistung im logischen Denken infolge der Art des visuellen Primings unterscheidet, wurde eine einfaktorielles Varianzanalyse (ANOVA) gemacht, obwohl eine der Voraussetzungen in Form der Normalverteilung der Leistungen nicht gege-

ben war. Der Kolmogorov-Smirnov-Test auf Normalität (Field, 2007) ergab hier eine signifikante Abweichung von der Normalverteilung ($D = .255$, $p = .000$). Da die global geprimte Gruppe ($n = 11$), die lokal geprimte Gruppe ($n = 18$) und die Kontrollgruppe ($n = 8$) zu klein waren für die Anwendung des Zentralen Grenzwerttheorems (Bortz 1999), wird die Verteilung im Diskussionsteil noch einmal aufgegriffen. Jedenfalls ergab die ANOVA keine signifikanten Unterschiede zwischen den Gruppen, $F(2, 34) = 0.771$, $p = .471$ (siehe Abbildung 2).

Für den Nachweis eines Einflusses des visuellen Primings wurden die Interrater-Reliabilität bestimmt und eine ANOVA vorgenommen. Cronbach's Alpha, das Förster & Denzler (2012) als Maß für die Interrater-Reliabilität verwendeten, war mit .83 vergleichbar mit der Originalstudie ($\alpha = .85$). Allerdings waren die Einstufungen der Rater in der Replikation homogener als reliabel: Krippendorff's Alpha ergab mäßige .47. Die ANOVA, für deren Anwendung die Voraussetzungen erfüllt waren, ergab wiederum keine signifikanten Unterschiede zwischen den Gruppen, $F(2, 34) = 0.149$, $p = .862$ (siehe Abbildung 3).

Somit beeinflusste den einfaktorischen ANOVAs zufolge das Priming in der Stichprobe weder das logische noch das kreative Denkvermögen.

Zusätzlich wurde zur Bestimmung eines Interaktionseffektes zwischen den Primingarten auf das Denkvermögen eine zweifaktorielle (logisch, kreativ) \times (global, lokal, KG) ANOVA angestellt. Dazu wurden die Leistungen der Versuchspersonen z-transformiert. Im Unterschied zu Förster & Denzler (2012) konnte kein Interaktionseffekt festgestellt werden, $F(2, 68) = 0.000$, $p = 1.000$, d.h. die verschiedenen Primings wirkten auf das logische und kreative Denkvermögen gleichermaßen – nämlich gar nicht.

Die Mittelwerte für die logischen und kreativen Leistungen liegen augenscheinlich nicht auf einer Geraden (siehe Abbildungen 4 und 5). Aus Vergleichsgründen wurden die Leistungen einer Linearen Regression und einem F-Test auf Linearität unterzogen. Die Prüfgröße

$$\Delta F = \frac{QS_{\text{regression}} - QS_{\text{zwischen}}}{MQ_{\text{innerhalb}}},$$

die bei einem linearen Zusammenhang gegen Null strebt (Anonymous, 2012), ergab für das logische Denkvermögen $\Delta F = 0.561$, $p(\Delta F) = .326$ (Förster: $\Delta F = 0.000$, $p(\Delta F) = 1.000$) und $\Delta F = 0.195$, $p(\Delta F) = .747$ (Förster: $\Delta F = 0.004$, $p(\Delta F) = .951$) für das kreative Denkvermögen.

Im Einklang mit Förster & Denzler (2012) konnte kein Einfluss der Stimmung auf das Denkvermögen nachgewiesen werden. Die einfaktorielle ANOVA der wegen der kleinen Zellengrößen klassierten Stimmungsausprägungen ergab auf die Leistung weder des logischen ($F(1,35) = 0.037$; $p = .849$ für Positive Affektivität und $F(1,35) = 0.409$; $p = .527$ für Negative Affektivität) noch des kreativen Denkens ($F(1,35) = 0.982$; $p = .328$ für Positive Affektivität und $F(1,35) = 0.220$; $p = .642$ für Negative Affektivität) einen signifikanten Effekt.

Während der Untersuchung änderte sich die Stimmung der Versuchspersonen nicht. Der t-Test für Positive Affektivität ($t(36) = 1.61$; $p = .116$) sowie für Negative Affektivität ($t(36) = -1.26$; $p = .216$) ergab keine signifikante Veränderung der Stimmung.

Diskussion

Förster und Denzler (2012) fanden, dass ein global geprägter Verarbeitungsstil die Leistung kreativen Denkens steigert und die logischen Denkens verschlechtert, wohingegen ein lokal geprägter Verarbeitungsstil die Leistung des logischen Denkens steigert und die kreativen Denkens verschlechtert. Dieser Befund konnte für visuell geprägte Verarbeitungsstile nicht repliziert werden. Weder die von Förster und Denzler (2012) beschriebenen Effekte noch gegenläufige Effekte stellten sich ein. Stattdessen konnte überhaupt kein Effekt des Primings auf die Leistung des kreativen oder logischen Denkens nachgewiesen werden.

Statt auf Effekte aus einer differenziellen Verarbeitung bei der Lösung kreativer und logischer Aufgaben deuten die Ergebnisse der Replikationsstudie vielmehr auf eine einheitliche Verarbeitung hin. Für den Modus einer einheitlichen Verarbeitung spricht das Phänomen des globalen Vorrangs, das besagt, dass Figuren schneller erkannt werden als die Elemente, aus denen sie zusammengesetzt sind (Navon, 1977). Beim globalen Vorrang handelt es sich um eine dispositionale Grundeinstellung der Informationsverarbeitung, die äußerst robust ist gegenüber Störungen: weder eine inkongruente oder heterogene Zusammensetzung der Navon-Letters (Navon, 2003) noch der explizite Hinweis, sich auf dessen Elemente (kleine Großbuchstaben) zu konzentrieren (Mills & Dodd, 2014), konnten den globalen Vorrang außer Kraft setzen. Soll es zu einer lokalen Verarbeitung kommen, muss die globale Verarbeitung erst gehemmt werden (Poirel et al., 2014). Bestimmte Figuren kann man zwar lernen zu ignorieren, doch die Disposition zum Vorrang der globalen Verarbeitung bleibt intakt (Navon, 2003).

Dem globalen Vorrang zufolge ist die Informationsverarbeitung eine globale Verarbeitung, die von lokalen Verarbeitungen unterbrochen ist. Demgemäß würden die Versuchspersonen einheitlich beim lokalen Priming kurz die kleinen Großbuchstaben identifizieren, um dann wieder in die globale Grundeinstellung der Verarbeitung zurückzukehren. Bei einer voreingestellten globalen Verarbeitung würde sich zudem ein globales Priming erübrigen. Die Replikationsstudie legt damit konsistent nahe, dass sich alle Priminggruppen im gleichen Verarbeitungsmodus befanden, einschließlich der Kontrollgruppe. Letztere nimmt in Försters Versuchsdesign eine bemerkenswerte Leerstelle ein, weil in seinem Verarbeitungsmodell GLOMO^{sys} kein neutraler Verarbeitungsstil vorkommt. Angesichts des globalen Vorrangs müssten sowohl die Kontrollgruppe als auch die global geprimte Gruppe Informationen global verarbeiten – und letztlich bei den Aufgaben vergleichbar abschneiden, wie in der Replikationsstudie. Dass die Leistungen der Kontrollgruppe bei Förster & Denzler (2012) in der Mitte, und zwar ziemlich genau in der Mitte zwischen der global und der lokal geprimten Gruppe lagen, ist in ihrer Linearität nicht nur statistisch, sondern auch theoretisch ungereimt. Erklären ließe sich das bestenfalls mit stetigen Intensitätsgraden des globalen Verarbeitungsstils. Davon ist aber in der Literatur nirgendwo die Rede.

Unterstellt man entgegen des globalen Vorrangs, dass es mittels Navon-Letters gelingt, entweder einen globalen oder einen lokalen Verarbeitungsstil anhaltend auszulösen, ist der ausgebliebene Interaktionseffekt der Replikationsstudie ein Hinweis darauf, dass die Verarbeitungsstile gleichermaßen an der Lösung kreativer wie logischer Aufgaben beteiligt sind und die Aufgabentypen insofern sich nicht grundsätzlich unterscheiden. In diesem Sinne ist Kreativität konzipiert worden als Erweiterung der allgemeinen Intelligenz (Guilford, 1959). Und in der Tat korrelieren die Leistungen bei den logischen Aufgaben des GRE-Tests, die von Förster und in der Replikationsstudie verwendet wurden, positiv – wenn auch nur leicht – mit der Kreativität der Versuchspersonen (Powers & Kaufman, 2004). Gegenläufige Leistungen wie bei Förster & Denzler (2012) wären demnach erst gar nicht zu erwarten gewesen.

Methodisch leidet das Versuchsdesign von Förster & Denzler (2012) daran, dass nur die Stimmung als Störvariable kontrolliert wird. Dabei wird schon allein der globale Vorrang bei der visuellen Wahrnehmung von mehreren Faktoren moderiert, wie von der Anzahl der Elemente oder der Vertrautheit mit ihnen – jemand, der mit hebräischen statt mit lateinischen Buchstaben vertraut ist, wird die lateinischen Buchstaben, wenn überhaupt, nur langsam identifizieren können; der Einfallswinkel auf die Retina spielt genauso eine Rolle, wie der Wahr-

nehmungswinkel (Im Wald sieht man eher den Baum als den Wald!) und die Wahrnehmungsdauer (Navon, 2003). Daher sollten in künftigen Studien zu Verarbeitungsstilen die Moderatorvariablen des globalen Vorrangs kontrolliert und insbesondere die Zeiten bis zur Identifikation von Figur oder Element gemessen werden. Letzteres würde zur Klärung beitragen, ob die Verarbeitung seriell (Crewther et al., 2007) oder parallel (Nakatani & Yamaguchi, 2014) zu modellieren ist.

Ebenfalls in künftigen Studien nach Möglichkeit zu kontrollierende Variablen sind Faktoren, die das kreative oder logische Denkvermögen direkt beeinflussen sollen, wie soziale Distanz (Liberman & Trope, 2008) und Machtposition (Smith & Trope, 2006) oder Konfession (Colzato, Wildenberg & Hommel, 2008) und Kulturzugehörigkeit (Kühnen & Oyserman, 2002). Dazu zählen auch Farben. Die Evidenzen sind hier jedoch widersprüchlich. Während Rot das logische Denkvermögen verbessern soll (Elliot & Maier, 2007), soll Grün das kreative Denkvermögen verbessern (Lichtenfeld, Elliot & Maier, 2012). Eine Verbesserung des kreativen Denkvermögens sollte auch Blau leisten (Mehta & Zhu, 2009), doch konnte diese Studie nicht repliziert werden (Steele, 2014). Eine Kontrolle dieser Variablen könnte helfen, die Widersprüche auszuräumen und zu klären, in welchem Verhältnis die beiden Denkvermögen zueinander stehen.

Da sich zudem Hinweise darauf finden, dass der Verarbeitungsstil beeinflusst wird von der Intelligenz und Persönlichkeitsmerkmalen, so dass es den Versuchspersonen verschieden leicht fällt, die lokalen Elemente gegen den globalen Vorrang zu identifizieren (Bouvet et al., 2011), ist eine Individualisierung künftiger Studien überlegenswert. Um die relativen Unterschiede eines Individuums in der kreativen und logischen Leistung zu bestimmen, müssten die Versuchspersonen vor der jeweiligen Bearbeitung der Aufgaben einmal lokal und zu einem anderen Zeitpunkt global geprimt werden. Dann sollten allerdings mehr Items als nur ein Cartoon verwendet werden, um bei der Beurteilung der kreativen Leistung zu valideren Werten zu gelangen (Csikszentmihalyi, 1997). So ließe sich differenzierter beurteilen, ob die Verarbeitungsstile, wie in der Replikationsstudie, sich gleichermaßen auf das kreative (De Dreu, Baas & Giacometonio, 2010) oder das logische Denkvermögen auswirken.

All die theoretischen und methodischen Schwächen von Förster & Denzler (2012) sollen nicht darüber hinwegtäuschen, dass auch die Replikationsstudie ihre Schwächen hat. Zuvörderst der geringe Stichprobenumfang von $N = 37$ und die damit verbundene geringe Teststärke (Tversky, 1971). Der Umfang liegt deutlich unter den anvisierten 60 bis 66

Versuchspersonen, anhand derer nur große Effekte hätten nachgewiesen werden können. Die konservative Herangehensweise wurde gewählt, weil in der Originalstudie große Effektstärken berichtet wurden. Es ist also durchaus möglich, dass die Verarbeitungsstile auf das Denkvermögen gegenläufige Effekte haben, die kleiner sind, als sie mit der Teststärke der Replikationsstudie nachgewiesen werden konnten. In der Replikationsstudie wären nur Effekte feststellbar gewesen, die noch größer hätten ausfallen müssen als in der Originalstudie (Simonsohn, 2013).

Dem Replikationsrezept (Brandt et al., 2014) folgend war die Studie nicht nur thematisch an Förster und Denzler (2012) angelehnt, sondern als direkte Replikation (Pashler & Harris, 2012) des visuellen Teils konzipiert. Dennoch könnten technisch bedingte Verschiedenheiten zwischen den Studien die Ergebnisse der Replikationsstudie verzerrt haben. Dies betrifft in erster Linie die Umsetzung des Primings und die Schwierigkeit der logischen Aufgaben.

In der Replikationsstudie folgte der nächste Primingreiz nicht auf Tastendruck, sondern auf zwei Mausklicks: erst Entscheidung zwischen L und H, danach auf 'Weiter'-Button. Das ist einerseits ermüdend, weshalb 34 Teilnehmer das Priming abgebrochen haben; bedeutsamer ist aber, dass aufgrund der Maussteuerung die Aufmerksamkeit der Versuchspersonen nicht allein den Buchstaben im oberen Drittel des Bildschirms galt, sie mussten darüber hinaus den Mauszeiger präzise auf die richtige Schaltfläche darunter führen, sich also auf Elemente der Bildschirmanzeige konzentrieren. Dadurch könnte ein lokales Priming ausgelöst worden sein, das das Priming in der globalen Gruppe konterkariert haben könnte.

Zusätzlich stellt den Erfolg des Primings der Umstand in Frage, dass im online-Setting der Replikationsstudie visuell außer der Zusammensetzung der Buchstaben keine Reize kontrolliert werden konnten, die gemäß GLOMO^{sys} den Verarbeitungsstil beeinflussen, wie beispielsweise Farben, so dass das Priming nicht allein auf die Buchstaben zurückging, wenn am Monitor einer Versuchsperson ein roter Aufkleber klebte oder daneben eine grüne Topfpflanze stand, wenn auf dem Schreibtisch ein rotes Schulheft lag, das Zimmer grün tapeziert war oder ein rotes Poster an der Wand hing. Ganz abgesehen davon, dass Försters Modell den Umgang mit inkongruenten Reizen – ein gutgelaunter südamerikanischer Bischof, der in einer norwegischen Holzkirche ganz nah vor einem roten Marienbild kniet – offen lässt, scheint es allerdings überhaupt fragwürdig, dass beim multimodalen Erleben einer Situation der Reiz einer Modalität allein den ausschließlichen Verarbeitungsstil der Wahrnehmung und der Begriffe festlegen soll (Rumelhart & McClelland 1986).

Der Verarbeitungsstil der Begriffe beim logischen Denken ist unerheblich, wenn der Schwierigkeitsgrad der Aufgaben so hoch ist, dass sie in der vorgegebenen Bearbeitungszeit und ohne Hilfsmittel nicht gelöst werden können. Dann bliebe der Effekt selbst eines erfolgreichen lokalen Primings unterschwellig und die Varianz wäre so weit eingeschränkt, dass anhand der Leistungen keine Differenzierung möglich ist. Dieser Bodeneffekt könnte in der Replikationsstudie eingetreten sein. Dafür spricht, dass alle Gruppen im Mittel nur eine der vier Aufgaben im Multiple Choice-Format gelöst haben, was bei jeweils fünf Antwortvorschlägen dem Erwartungswert für bloßes Raten sehr nahe kommt. Raten war für die Versuchspersonen verlockend, weil logisches Denken anstrengend ist und grundsätzlich vermieden wird (Kahneman 2012) und weil Anreize fehlten, sich der Anstrengung zu stellen: sie erhielten weder eine Entschädigung noch konnten sie sich wegen der Anonymisierung durch eine überdurchschnittliche Leistung hervortun. Entsprechend wäre eine Gleichverteilung der Resultate zu erwarten, was zudem den Mangel ihrer Normalverteilung erklären würde. Dagegen spricht allerdings, dass die erste (40 %) und die zweite Aufgabe (24 %) insgesamt deutlich häufiger gelöst wurden als die dritte und vierte Aufgabe (jeweils 8 %). Hier scheint die Zeit nur für eine ernsthafte Bearbeitung der ersten Aufgabe gereicht zu haben. Da andere Einflussgrößen, wie Intelligenz oder Konzentrationsvermögen, durch die Randomisierung der Gruppen neutralisiert sein sollten, ist nicht von einem Raten auszugehen. Insofern spricht vieles für einen Bodeneffekt, der die Aussagekraft der Replikationsstudie zum logischen Denken erheblich schwächt.

Die Aussagekraft zum kreativen Denken leidet an der geringen Interrater-Reliabilität. Geringe Reliabilitäten sind in der Kreativitätsforschung jedoch nicht ungewöhnlich (Kim, 2006). Hier harren die Gütekriterien noch einer adäquaten Konzeptualisierung, anhand derer die Rater im Vorfeld hätten geschult werden können, um zu reliableren Ergebnissen zu gelangen.

Als Reliabilität wurden bei 18 bis 29jährigen für die mit der Kurzversion des PANAS ermittelte Stimmung hohe interne Konsistenzen (Cronbach's Alpha) berichtet (Mackinnon et al., 1999): $\alpha = .75$ für Positive Affektivität und $\alpha = .86$ für Negative Affektivität. Bemerkenswert an den Ergebnissen zur Stimmung ist nur, dass Förster und Dannenberg (2010) im GLOMO^{sys} wie Davis (2009) oder Fernández-Abascal & Martin (2013) eine differenzielle Wirkung der Stimmung auf Verarbeitungsstil und kognitive Leistung postulieren, Förster &

Denzler (2012) dagegen lapidar das Fehlen einer solchen Wirkung feststellen, ohne diesen widersprüchlichen Befund zu diskutieren.

Solche Ungereimtheiten im Modell machen einen stutzig. Allen voran die ominöse Kontrollgruppe, die in GLOMO^{sys} gar nicht abgebildet ist, deren Leistung allerdings so genau in die Mitte passt zwischen die Leistung der globalen und der lokalen Gruppe, dass das Zustandekommen eines solchen linearen Zusammenhangs statistisch nahezu ausgeschlossen ist (Anonymous, 2012). In dieses Bild passt, dass Förster & Denzler (2012) in ihren Diagrammen zur kreativen und logischen Leistung als Streuungsmaß den Standardfehler angeben. (In den Abbildungen 2 und 3 ist dies aus Gründen der Vergleichbarkeit erfolgt.) Im Unterschied zur Standardabweichung verrät der Standardfehler nichts von der Stichprobe, fällt aber kleiner aus als jener und täuscht so auf den ersten Blick eine geringere Streuung vor.

Die Verwendung eines Schätzwertes zur Angabe der Streuung mag man als legitime Kosmetik durchgehen lassen. Anders verhält es sich dagegen mit handwerklichen Fehlern wie dem folgenden. Keinesfalls stimmen können nämlich die Freiheitsgrade, die Förster & Denzler (2012) für den Interaktionseffekt berichten. Denn die Versuchspersonen gehen dort zweimal ein in die Varianzanalyse: einmal mit ihrer logischen und einmal mit ihrer kreativen Leistung. Und selbst wenn die Autoren die Analyse an 60 Werten vorgenommen hätten, bspw. durch Addition der z-Werte zu einer Gesamtsumme, ergäbe sich daraus ein $F(2, 54)$ statt des angegebenen $F(2, 57)$. So gewinnt man den Eindruck, dass Förster entweder nicht weiß, was er tut, oder eben die Zahlen einfach produziert.

Dagegen nimmt sich das L aus Hs, das Förster und Denzler (2012) beim Priming verwendet haben wollen, das dafür aber ganz offensichtlich nicht geeignet ist, wie ein Flüchtigkeitsfehler aus. Im Kontext all der anderen genannten Auffälligkeiten gewinnt er aber an Schwere. Zusammengenommen kratzen sie empfindlich an der Integrität von Försters Forschung. Diesbezüglich wiegt der Umstand am geringsten, dass ein Teil seiner Studie hier nicht repliziert werden konnte. Es wurde eine Vielzahl von Gründen dafür angeführt, weshalb die Replikation nicht geglückt sein könnte. Diese haben nichts zu tun mit einer Datenmanipulation. Der Datenmanipulation kann man einen Forscher im Grunde nur überführen, wenn man ihn in flagranti erwischt (Stroebe, Postmes & Spears, 2012).

Genauso wenig hätte eine erfolgreiche Replikation Förster entlastet. Einen Fälschungsverdacht kann eine Replikation nur insofern aufklären helfen, als dafür die zu replizierende Studie sehr genau analysiert werden muss. Die Analyse von Förster und Denzler (2012)

förderte Lässlichkeiten, Fehler und Ungereimtheiten zutage, die sich über Förster verdichten zu einer schwarzen Gewitterwolke. Und es sieht nicht so aus, als würden sie bald vorüberziehen.

Literatur

- Anonymous. (2012). Suspicion of scientific misconduct by Dr. Jens Förster. https://retractionwatch.files.wordpress.com/2014/04/report_foerster.pdf. (Zugriff am 30.12.2014)
- Asendorpf, J.B. et al. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27(2), 108–119.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. (5. Aufl.) Berlin: Springer.
- Bouvet, L., Rousset, S, Vaklois, S. & Donnadieu, S. (2011). Global precedence effect in audition and vision: Evidence for similar cognitive styles across modalities. *Acta Psychologica*, 138 (2), 329-335.
- Brandt, M. J. et al. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217-224.
- Colzato, L. S., van den Wildenberg, W., & Hommel, B. (2008). Losing the big picture: How religion may control visual attention. *PLoS ONE*. doi: 10.1371/journal.pone.0003679.
- Crewther, D. P. et al. (2007). Global and local attention in the attentional blink. *Journal of Vision*, 7(14), 1-12.
- Csikszentmihalyi, M. (1997). *Creativity*. New York: HarperCollins.
- Davis, M. A. (2009). Understanding the relationship between mood and creativity: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 108, 25–38.
- De Dreu, C. K., Baas, M., & Giacomettonio, M. (2010). Processing modes and creativity: Why (not)? *Psychological Inquiry*, 21(3), 203-208.
- Elliot, A. J., & Maier, M. A. (2007). Color and psychological functioning. *Current Directions in Psychological Science*, 16, 250-254.
- Faul, F. et al. (2007). G*Power3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Fernández-Abascal, E. & Martín, M. (2013). Affective induction and creative thinking. *Creativity Research Journal*, 25(2), 213-221.
- Field, A. (2005). *Discovering Statistics Using SPSS* (2. Aufl.). London: Sage.
- Förster, J. (2009). Relations between perceptual and conceptual scope: How global versus local processing fits a focus on similarity versus dissimilarity. *Journal of Experimental Psychology: General*, 138, 88–111.
- Förster, J. & Dannenberg, L. (2010). GLOMO^{sys}: A systems account of global versus local processing. *Psychological Inquiry*, 21, 175–197.
- Förster, J. (2011). Local and Global Cross-Modal Influences Between Vision and Hearing, Tasting, Smelling, or Touching. *Journal of Experimental Psychology: General*, 140, 364-389.

- Förster, J. & Denzler, M. (2012). Sense creative! The impact of global and local vision, hearing, touching, tasting and smelling on creative and analytic thought. *Social Psychological and Personality Science*, 3(1), 108-117.
- Friedman, R. & Förster, J. (2000). The effects of approach and avoidance motor actions on the elements of creative insight. *Journal of Personality and Social Psychology*, 79, 477-492.
- Friedman, R. & Förster, J. (2002). The influence of approach and avoidance motor actions on creative cognition. *Journal of Experimental Social Psychology*, 38, 41-55.
- Gasper, K. & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science*, 13, 33-39.
- Gilovich, T., Vallone, R., Tversky, A. (1985) The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17, 295-314.
- Guilford, J.P. (1959). *Personality*. New York: McGraw Hill.
- Kahneman, D. (2012). *Thinking, fast and slow*. London: Penguin.
- Kim, K.H. (2006) Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal* 18(1), 3-14.
- Kolfschooten, F.v. (2014). Psychologist's defense challenged. *Science*, 344, 957-958.
- Krohne, H. W., Egloff, B., Kohlmann, C.-W. & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der «Positive and Negative Affect Schedule» (PANAS). *Diagnostica*, 42, 139-156.
- Kühnen, U., & Oyserman, D. (2002). Thinking about the self influences thinking in general: Procedural consequences of self-construal activation. *Journal of Experimental Social Psychology*, 38, 492-499.
- Lieberman, N., & Trope, Y. (2008). The psychology of transcending the here and now. *Science*, 322, 1201-1205.
- Lichtenfeld, S., Elliot, A.J., Maier, M.A., Pekrun, R. (2012). Fertile green: Green facilitates creative performance. *Personality and Social Psychology Bulletin*, 38(6), 784-797.
- Mackinnon, A., Jorm, A.F., Christensen, H., Korten, A.E., Jacomb, P.A., & Rodgers, B. (1999). A short form of the positive and negative affect schedule: evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences*, 27, 405-416.
- Maier, M. A., Elliot, A. J., Lichtenfeld, S.. (2008) Mediation of the negative effect of red on intellectual performance. *Personality and Social Psychology Bulletin*, 34, 1530-1540.
- Mehta, R. & Zhu, R. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, 323, 1226-1229.
- Mills, M. & Dodd, M. D. (2014). Which way is which? Examining global/local processing with symbolic cues. *Journal of Experimental Psychology: General*, 143(4), 1429-1436.

- Nakatani, H. & Yamaguchi, Y. (2014). Quick concurrent responses to global and local cognitive information underlie intuitive understanding in board-game experts. *Scientific Reports*, 4, 5894. doi:10.1038/srep05894.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica*, 114, 273-309.
- Pashler, H. & Harris, C.R. (2012). Is the replicability crisis overblown? *Perspectives on Psychological Science*, 7, 531-536.
- Poirel, N. et al. (2014). Do you want to see the tree? Ignore the forest: inhibitory control during local processing: a negative priming study of local-global processing. *Experimental Psychology*, 61 (3), 205-214.
- Powers, D.E. & Kaufman J. C. (2004) Do standardized tests penalize deep-thinking, creative, or conscientious students? Some personality correlates of Graduate Record Examinations test scores. *Intelligence*, 32 (2), 145-153.
- Rumelhart, D.E. & McClelland, J.L. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Cambridge: MIT Press.
- Simonsohn, U. (2013). Evaluating replication results. <http://ssrn.com/abstract=2259879>. (Zugriff am 08.02.2015)
- Smith, P. K., & Trope, Y. (2006). You focus on the forest when you're in charge of the trees: Power priming and abstract information processing. *Journal of Personality and Social Psychology*, 90, 578-596.
- Steele, K.M. (2014) Failure to replicate the Mehta and Zhu (2009) color-priming effect on anagram solution times. *Psychonomic Bulletin & Review*, 21(3), 771-776.
- Stroebe, W., Postmes, T. & Spears, R. (2012) Scientific misconduct and the myth of self-correction in science. *Perspectives of Psychological Science* 7(6), 670-688.
- Tversky, A. & Kahneman, D. (1971) Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105-110.
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54, 1063-1070.
- Wronska, M. & Blazejewicz, B. (2013). Non-affective cues influence problem solving. <http://psychfiledrawer.org/replication.php?attempt=MTc4>. (Zugriff am 30.12.2014)

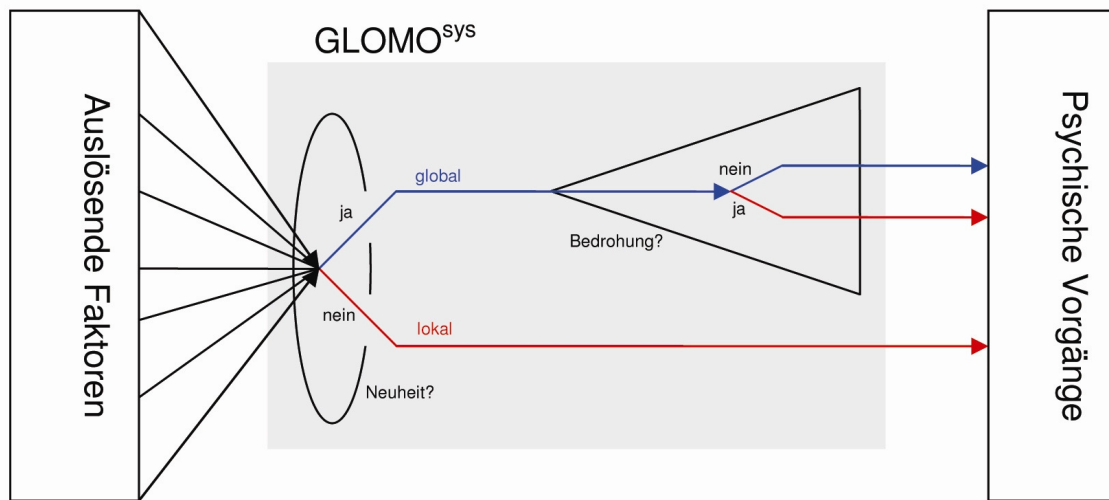


Abbildung 1. Eigene schematische Darstellung von GLOMO^{sys} nach Förster & Dannenberg (2010).

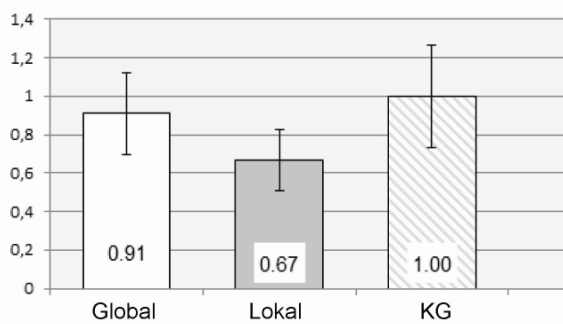


Abbildung 2. Durchschnittliche Anzahl der richtigen Antworten beim logischen Denken. Die Zahlen repräsentieren die Mittelwerte. Die Fehlerbalken den Standardfehler.

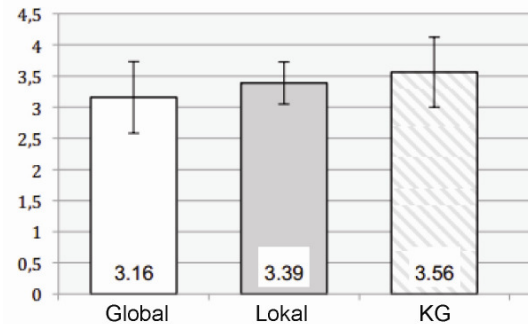


Abbildung 3. Durchschnittliche Bewertung der Kreativität. Die Zahlen repräsentieren die Mittelwerte. Die Fehlerbalken den Standardfehler.

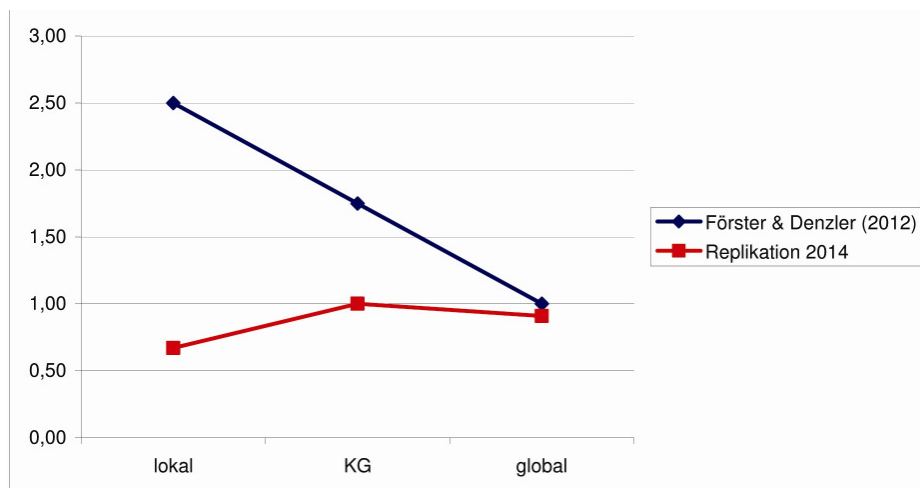


Abbildung 4. Durchschnittliche Anzahl der richtigen Antworten beim logischen Denken mit Mittelstellung der Kontrollgruppe (KG) im Vergleich.

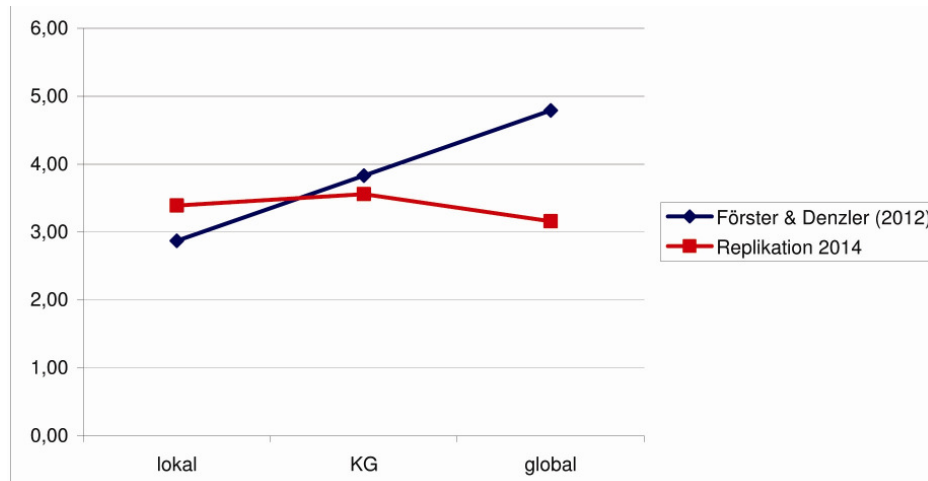


Abbildung 5. Durchschnittliche Bewertung der Kreativität mit Mittelstellung der Kontrollgruppe (KG) im Vergleich.