

Kurt Haim
 Wolfgang Aschauer
 Christoph Weber

Pädagogische Hochschule OÖ

Diagnose divergenter Denkstrategien bei naturwiss. Problemstellungen

Theoretischer Hintergrund

Problemlösen kann als zielorientiertes Denken und Handeln in Situationen definiert werden, für deren Bewältigung keine routinierten Vorgehensweisen verfügbar sind (Klieme et al., 2001). Problemlösen erfolgt also, wenn ein unerwünschter Ausgangszustand in einen erwünschten Endzustand transformiert werden soll, diese Transformation jedoch durch Barrieren behindert wird (OECD, 2013). Problemlösen ist daher ein komplexer Prozess, bei dem Faktoren wie Persönlichkeit, Motivation Umfeld und divergente Denkstrategien eine Rolle spielen (Meador, 1997).

Divergentes Denken (DT für „Divergent Thinking“) beinhaltet kognitive Prozesse, die über das Wechseln von Perspektiven zum Finden mehrerer Antworten auf offene Fragen oder Probleme führt (Kaufman et al., 2008). Mithilfe DT können bei Problemstellungen verschiedenste potenzielle Lösungen abgeleitet werden, was statistisch gesehen eine hohe Wahrscheinlichkeit für eine erfolgreiche Problemlösung nahelegt (Kaufman et al., 2008). Daher stellt divergentes Denken in diesem Zusammenhang einen besonders wichtigen Indikator für kreatives Problemlösepotential dar (Runco & Agar, 2012).

In der Literatur werden häufig drei Aspekte divergenten Denkens erwähnt: Fluidität, Flexibilität und Originalität. Fluidität ist die Anzahl der Antworten auf ein bestimmtes Problem. Flexibilität ist die Anzahl der Kategorien aus denen die Ideen stammen und Originalität ist die Einzigartigkeit einer genannten Lösung (Runco, 1999).

Viele Testverfahren für DT untersuchen vor allem Fluidität und Flexibilität (Runco & Acar, 2012). Der international am meisten verwendete Test für divergentes Denken ist der Torrance Test (Torrance, 1966). Er beinhaltet figurale Aufgaben, wie die Konstruktion von Bildern als auch verbale Aufgaben, bei denen nach Konsequenzen, Verwendungen, etc. gefragt wird. International kommt auch der „Divergent Thinking-Test“ von Runco (Chand & Runco, 1993) zum Einsatz. Hier findet man Aufgaben, in denen viele Verwendungen von bestimmten Gegenständen gefragt sind („Many Uses“) sowie Aufgaben, in denen unterschiedlichste Lösungen von Alltagsproblemen zu generieren sind („Realistic Problems“).

So valide diese Tests auch sind, weisen sie aus Sicht der Autoren auch Defizite auf. So erfolgt die Auswertung aller Antworten oft mit einem allgemeinen Kategoriensystem, was die Auswertung für die jeweiligen Items oft erheblich erschwert. Für viele Items fehlt oft ein Kategoriensystem, was eine Bewertung hinsichtlich Flexibilität unmöglich macht. Weiters findet man in den oben genannten Tests selten Items mit einer klaren Domänenspezifität hinsichtlich naturwissenschaftlicher Probleme.

Für die Erhebung divergenten Denkens bei naturwissenschaftlichen Problemstellungen entwickelten daher die Autoren ein eigenes Diagnoseinstrument, den ADTS-Test (Assesment of Divergent Thinking in Science). Der ADTS-Test enthält 6 Items mit naturwissenschaftlichem Fokus samt Kategoriensystem für jedes einzelne Item. Dieser Test ist als Online-Test konzipiert und sieht eine Bearbeitungszeit von 3 Minuten pro Item vor. Für die Auswertung, die durch zwei Rater erfolgte, wurden Fluidität und Flexibilität erhoben und daraus der Kreativitätsquotient (CQ) berechnet (Snyder et al., 2004):

$$CQ = {}^2\log\{(1 + u_1) \cdot (1 + u_2) \cdot \dots \cdot (1 + u_c)\}$$

$u_i \dots$ Zahl der Nennungen der i -ten Kategorie ($i = 1, 2, \dots, c$)

Der CQ liefert eine Maßzahl, in der Fluidität und Flexibilität zusammengefasst werden, wobei die Flexibilität höher gewichtet wird.

Forschungsziele

Zur Förderung des divergenten Denkens bei Jugendlichen entwickelten die Autoren spezielle Unterrichtstechniken für den Chemie-, Physik- und Biologieunterricht, die mit dem Begriff „flex-based learning“ zusammengefasst werden. Um die Wirksamkeit der entwickelten Unterrichtstechniken im Rahmen einer Interventionsstudie messen zu können, wurde hierfür der ADTS entwickelt. Ziel der hier beschriebenen Arbeit ist die Validierung sowie die Reliabilitätsprüfung des ADTS.

Forschungsdesign

Im Rahmen der Studie wurden insgesamt 1103 Schülerinnen und Schüler getestet. Die Probanden stammten aus 51 Klassen aus 35 österreichischen Schulen. Im September 2018 erfolgte ein Prätest. Dieser beinhaltete 4 Items des ADTS und 3 Items des Runco-Tests, wobei die Probanden randomisiert entweder 3 Items des Subtests *Many Uses* (MU) oder 3 Items *Realistic Problems* (RP) erhielten.

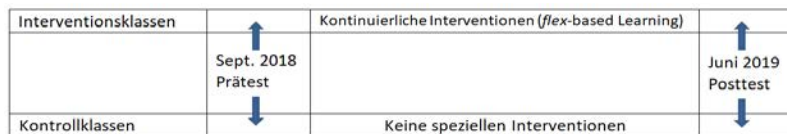


Abb. 1: Grafische Darstellung des Forschungsdesigns

In den Kontrollklassen (318 Probanden) wurde im gesamten Schuljahr keine spezielle Intervention durchgeführt. In den Interventionsklassen (785 Probanden) wurden kontinuierlich während des gesamten Schuljahrs Unterrichtstechniken des *flex-based learnings* eingesetzt. Die mittlere Anzahl an Interventionen betrug dabei 10. Im Juni 2019 (Posttest) wurden alle Schülerinnen und Schüler nochmals getestet. Dabei wurden nur Items des ADTS verwendet, wobei drei Items vom Prätest und zwei neue Items zum Einsatz kamen.

Ergebnisse

Die Interrater-Reliabilität wurde sowohl auf Ebene des Kategoriensystems als auch auf Ebene der latenten Variable CQ ermittelt. Für alle Items liegen die entsprechenden Werte in einem sehr zufriedenstellenden Bereich (siehe Tab. 1).

Tab. 1: Interrater-Reliabilität

Item	1	2	3	4	5	6
Kappa*	0.918	0.890	0.833	0.824	0.886	0.903
ICC**	0.993	0.954	0.975	0.974	0.972	0.984
* Cohen's Kappa auf Ebene des Kategoriensystems						
** Intra-Klassen-Korrelation (ICC) auf Ebene CQ						

Eine durchgeführte Hauptkomponentenanalyse sowie eine Reliabilitätsanalyse ergaben, dass die ADTS Skala eindimensional und intern konsistent ist. Die Alphawerte für den ersten und zweiten Messzeitpunkt liegen bei 0.782 bzw. 0.822. Die Korrelationen zwischen den einzelnen Items liegen in einem Bereich von 0.43 und 0.63.

Eine konfirmatorische Faktorenanalyse zeigte, dass alle Items signifikant ($p < .001$) und substantiell (Faktorladungen $> .63$) auf den latenten Faktor DT luden. Weiters kann auf eine akzeptable Passung zwischen dem theoretischen Messmodell und den empirischen Daten geschlossen werden, da alle Fit-Indizes CFI, RMSEA und SRMR sehr zufriedenstellend waren (CFI = .993, RMSEA = .06, SRMR = .015). Einzig die Chi² Statistik ist grenzwertig (Chi² = 9.782, df = 2), wobei dieser auf Grund der Anfälligkeit bei großen Stichproben keine zu große Bedeutung zugewiesen werden soll (vgl. Linder et al., 2019).

Die konvergente Validität des ADTS wurde mithilfe der Items des Runco-Tests erhoben, indem zwei Modelle (1 Faktor bzw. 2 Faktoren) gefittet wurden. Die Ergebnisse zeigen eine hohe Korrelation zwischen ADTS und Runco-Test, wobei wie erwartet die RP-Skala etwas höher korreliert als die MU-Skala. Trotz dieser Korrelation zeigen die Ergebnisse (siehe Abb. 2), dass das 2 Faktoren Modell besser fittet und der ADTS somit ein konzeptionell anderes Konstrukt abbildet.

	Chi ²	df	p	Delta Chi ² (df = 1)(a)	CFI	RMSEA	SRMR	r
<i>Realistic Problems</i>								
Model 1: 1 Faktor	67,7	14	0		0,989	0,06	0,034	
Model 2: 2 Faktoren	26,34	13	0,015	74,96***	0,997	0,031	0,021	0,876
<i>Many Uses</i>								
Model 1: 1 Faktor	158,688	14	0		0,984	0,098	0,055	
Model 2: 2 Faktoren	23,17	13	0,04	135,52***	0,999	0,027	0,02	0,754

Abb. 2: Auswertung bzgl. konvergenter Validität

Die zeitliche Invarianz des ADTS wurde anhand von drei Modellen (konfigural, metrisch und skalar) geprüft. Die Δ CFI-Werte liegen bei 0.001 für die metrische und bei 0.003 für die skalare Invarianz. Auf Basis dieser Δ CFI-Werte kann das Messmodell bzw. die ADTS-Skala als zeitlich invariant eingestuft werden und ist somit für Interventionsstudien geeignet.

<i>Konditionales Modell</i>	Mittelwert Coeff (SE)	Varianz Coeff (SE)	Kovarianz Coeff (SE)	Intervention Coeff (SE)
<i>Schülerebene (innerhalb von Klassen)</i>				
DT_T1		1,220*** (0,109)		
Delta_DT		1,010*** (0,145)	-0,302** (0,091)	
<i>Klassenebene (zwischen Klassen)</i>				
DT_T1	3,255 (0,219)	0,620*** (0,162)		0,066 (0,258)
Delta_DT	0,165 (0,102)	0,103* (0,045)	-0,046 (0,053)	0,441*** (0,136)
<i>Model Fit</i>				
	Chi ² (df)	CFI	RMSEA	SRMR
	212,759*** (68)	0,95	0,044	0,040/0,073

Abb. 3: Unterschiede auf Schüler- und Klassenebene und Einfluss der Intervention (Zweiebenen Second Order Latent Difference Modell)

Der Vergleich zwischen Prä- und Posttest zeigt, dass eine statistisch hochsignifikante Steigerung in der divergenten Denkfähigkeit (Delta_DT) auftritt, mit signifikanten Unterschieden auf Schüler- als auch auf Klassenebene. Die Unterschiede auf Klassenebene können dabei auf die Intervention zurückgeführt werden (siehe Abb. 3).

Diskussion

Zusammenfassend kann der ADTS als valides, reliables, Zeit invariantes und änderungssensitives Instrument zur Messung der domänenspezifischen divergenten Denkfähigkeit im naturwissenschaftlichen Kontext bezeichnet werden. Die hohe Interrater-Reliabilität spricht für das elaborierte Kategoriensystem. Erst dadurch kann neben der Fluidität auch die Flexibilität und somit der CQ bestimmt werden. Für die Validitätsprüfung wurde ein sehr einfaches Modell herangezogen, das Faktoren wie das Fachwissen, Alter und Schultype noch nicht berücksichtigt hat. In weiteren Analysen sollen die genannten Faktoren mit einbezogen und dadurch das theoretische Messmodell ausgebaut werden.

Die im Rahmen der Validierung erhobenen Daten werden in einem nächsten Schritt herangezogen, um den Interventionseffekt von flex-based learning messen zu können. Dabei werden zusätzlich Gendereffekte, Einfluss von Hochbegabung, Problemlösekompetz, etc. analysiert.

Literatur

- Case, R. (1977). Implications of developmental psychology for the design of effective instruction. In I.J.W. Pellegrino, E.S.D. Fokkema & R. Glaser (Eds.), *Cognitive psychology and instruction*. New York: Plenum, 441 – 465
- Chand, I. & Runco, M.A. (1993). Problem finding skills as components in the creative process. *Personality & Individual Differences*, 14, 155-162
- Duit, R., Gropengießer, H., & Stäudel, L. (2004). *Naturwissenschaftliches Arbeiten: Unterricht und Material 5-10*. Seelze-Velber: Friedrich-Verlag
- Kaufman, J.; Plucker, J.; Baer, J. (2008) *Essentials of Creativity Assessment*. Hoboken: Wiley
- Klieme, E.; Funke, J.; Leutner, D.; Reimann, P.; Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie. *Zeitschrift für Pädagogik* 47, 179-200
- Lederman, N.G., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R.S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching*, 39 (6), 497-521
- Lindner, C.; Lindner, M.A.; Retelsdorf, J. (2019). Die 5-Item-Skala zur Messung der momentan verfügbaren Selbstkontrollkapazität (SMS-5) im Lern- und Leistungskontext. *Diagnostica* 65 (4), 228–242
- Meador, K.S. (1997). *Creative thinking and problem solving for young learners*. Englewood: Teacher Ideas Press.
- OECD (2014). *PISA 2012 Results: Creative Problem Solving: Students' Skills in Tackling Real-Life Problems (Volume V)*, PISA, OECD Publishing.
- Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, 24(1), 66-75.
- Runco, M.A. (1999). Divergent thinking. In M.A. Runco & S. Pritzker (Eds), *Encyclopedia of creativity* (Vol. I.) San Diego: Academic Press, 577-582
- Torrance, E. P. (1966). *The Torrance Tests of Creative Thinking – Norms- Technical Manual Research Edition – Verbal Tests, Forms A and B – Figural Tests. Forms A and B*. Princeton NJ: Personnel Press