

## **Wirksamkeit von Feedback-Maßnahmen im Chemieunterricht der Sek. I**

### **Einleitung**

Die formative Beurteilung von Lernprozessen meint eine lernprozessbegleitende Diagnose des Kompetenzstandes und der Lernstrategien der Lernenden, die ein systematisches Feedback ermöglicht (Shepard, 2000; Cowie & Bell, 2001; OECD, 2005; Jürgens & Lissmann, 2015). Im Sinne einer individuellen Förderung kann die formative Beurteilung als zyklischer Prozess aus Kompetenzmessung, Diagnose, Rückmeldung und Zielsetzung verstanden werden, der durch die aktive Teilnahme der Lernenden an allen Schritten des Prozesses an Effektivität gewinnt (Stiggins & Chappuis, 2005). Insbesondere Rückmeldungen in Form eines Feedbacks scheinen einen starken Einfluss auf die Lernleistung zu haben (Hattie & Timperley, 2007; Rakoczy et al., 2008).

### **Theoretische Fundierung**

Im Jahr 2001 wurde von der Kultusministerkonferenz ein Maßnahmenkatalog verabschiedet, der die Implementation individueller Fördermaßnahmen in die Schulpraxis in Deutschland vorsah (KMK, 2002). Daran anschließend trat mit der Verankerung des Rechts jedes Lernenden auf eine individuelle Förderung in den Schulgesetzen der Länder ein einklagbares Grundrecht in Kraft, das die verpflichtende Förderung jedes Einzelnen entsprechend seiner Leistungsstärke, Kompetenzen, Fähigkeiten, Neigungen und weiteren individuellen Voraussetzungen vorsieht (Ministerium für Schule und Weiterbildung des Landes NRW, 2005). Im selben Jahr veröffentlichte die OECD einen Bericht, der zur flächendeckenden Implementation der formativen Beurteilung in den Schulen auffordert, um die Lernprozesse von Schülerinnen und Schülern individuell begleiten und fördern zu können (OECD, 2005). Durch die formative Beurteilung von Lernprozessen konnten positive Effekte auf die Motivation und Lern- und Leistungsentwicklung nachgewiesen werden (Stiggins, 2007). Weitere Studien zeigen, dass die formative Beurteilung die Selbstständigkeit und das Selbstwertgefühl der Lernenden steigern kann (Black & Wiliam, 1998). Als zentrales Element der formativen Beurteilung wird immer wieder das Feedback an die Lernenden genannt (McMillan, 2010; Wiliam & Leahy, 2007). Feedback scheint einen starken Einfluss auf die Lernleistung zu haben und gilt als essentielle Einflussgröße von effektiven Lehr- und Lernprozessen (Hattie & Timperley, 2007; Rakoczy et al., 2008). Vor diesem Hintergrund wird in dem Projekt eine Unterrichtseinheit für die Sekundarstufe I entwickelt und evaluiert, die auf Basis der formativen Beurteilung von Lernprozessen gestaltet ist und verschiedene Feedback-Maßnahmen enthält. Die Unterrichtseinheit behandelt den Themenbereich Atombau und beinhaltet unterschiedliche Feedback-Maßnahmen für drei Gruppen von Lernenden, um einen Vergleich von lehrergeleitetem und selbstreguliertem Feedback zu ermöglichen.

### **Forschungsfragen**

Im Rahmen der Unterrichtseinheit soll untersucht werden, ob die Durchführung der Einheit das unmittelbare und nachhaltige Fachwissen der Lernenden steigern kann und ob die Bearbeitung verschiedener Feedback-Maßnahmen zu einem unterschiedlichen Fachwissenszuwachs, zu einer unterschiedlichen Bearbeitungsqualität und zu einer unterschiedlichen Einstellung führt. Ebenso wird die Passung der verschiedenen Feedback-Maßnahmen für bestimmte kognitive Niveaus der Lernenden untersucht.

### **Ausgewählte Test- und Auswertungsinstrumente**

Das *Fachwissen* wird mit Hilfe eines Multiple-Choice-Tests (Cronbachs  $\alpha = .812$ ) erfasst. Dieser beinhaltet 25 Items mit einer korrekten Antwort von fünf (Holländer & Melle, 2012) und wird zu einem Zeitpunkt vor (Pre) und zu zwei Zeitpunkten nach (Post, Follow-up) der Intervention eingesetzt. Jeweils vor der Intervention werden die kognitiven Fähigkeiten mit dem CFT 20 (Weiß, 1998) und das schulische Selbstkonzept in den Fächern Chemie und Mathematik mit dem DISK-GITTER (Rost et al., 2007) erfasst. Die Einstellung zu der Unterrichtsreihe und zu den Feedback-Maßnahmen wird mit 43 Items (Subskala Unterrichtsreihe 31 Items, Subskala Vortrag 6 Items, Subskala Feedback-Maßnahmen 6 Items) nach der Intervention erhoben, wobei eine 5-stufige Likert-Skala von 1 = *stimme voll zu* (positiv) bis 5 = *stimme gar nicht zu* (negativ) verwendet wurde (Cronbachs  $\alpha_{\text{ges}} = .950$ , Cronbachs  $\alpha_{\text{Unterrichtsreihe}} = .945$ , Cronbachs  $\alpha_{\text{Vortrag}} = .806$ , Cronbachs  $\alpha_{\text{Feedbackmaterial}} = .819$ ). Weiterhin wurde ein Kodiermanual zur Untersuchung der Bearbeitungsqualität der Schülerarbeitsmaterialien ( $\text{ICC}_{\text{unjust}} = .971$ ) entwickelt. Dazu wurden alle von den Probanden bearbeiteten Feedback-Maßnahmen eingescannt und auf einer 4-stufigen Likert-Skala (von 1 = *nicht zufriedenstellend* bis 4 = *sehr gut*) bewertet. Zusätzlich wurde ein Kodiermanual zur Auswertung von Schüler- und Lehrerinterviews entwickelt, die nach der Intervention geführt wurden ( $\text{ICC}_{\text{unjust}} = .950$ ).

### **Erste Ergebnisse**

Die Datensätze sind noch nicht vollständig ausgewertet, sodass sich diese ersten Ergebnisse im Verlauf der Auswertung noch verändern können.

#### *Stichprobe*

Die Hauptuntersuchung wurde mit zehn 8. Klassen an Gymnasien durchgeführt, die Stichprobengröße beträgt  $N = 185$  (94 Mädchen und 91 Jungen). Für die Auswertung wurden die Daten derjenigen Probanden berücksichtigt, die sowohl an der Intervention als auch an der Pre-, Post- und Follow-up-Testung teilgenommen haben.

#### *Untersuchungsdesign*

Die Hauptuntersuchung wurde als experimentelles Design mit drei Interventionsgruppen durchgeführt. Die Interventionsgruppen bestehen aus Lernenden einer Klasse, die zur Gruppeneinteilung nach dem kognitiven Niveau und dem Vorwissen parallelisiert wurden. Zum ersten Messzeitpunkt werden die kognitiven Fähigkeiten, das schulische Selbstkonzept und das Fachwissen der Probanden erfasst. Danach erfolgt die Intervention mit sechs Unterrichtsstunden à 90 Minuten. Für die Durchführung der Intervention erhielten die Lehrkräfte Unterrichtsskizzen für jede Unterrichtseinheit und die Schülerarbeitsmaterialien. Während der Intervention wird ein Zyklus aus Beurteilung, Diagnose, Rückmeldung und Förderung mehrfach durchlaufen. Dazu findet am Ende einer Unterrichtseinheit eine 15-minütige Testphase statt, in der von den Lernenden aller Gruppen ein Kompetenztest bearbeitet wurde. Zu Beginn der nächsten Unterrichtseinheit erhielt die Feedbackgruppe (FB) einen Feedbackbogen mit individuellem, elaboriertem Feedback und die Lösungsbogengruppe (LB) einen Lösungsbogen, um sich selbstständig zu korrigieren. Die Lernenden der Ohne-Rückmeldungsgruppe (OR) haben die Möglichkeit die Bearbeitung der Aufgaben des Kompetenztests fortzusetzen. Die Dauer dieser Arbeitsphase beträgt zehn Minuten. Darauf folgt ein kurzer lehrerzentrierter Vortrag in Form einer Power-Point-Präsentation (fünf Minuten), der den Inhalt des Kompetenztests zusammenfasst. Danach arbeiten die Lernenden aller Gruppen gemeinsam an der Erarbeitung weiterer thematischer Inhalte und ein neuer Zyklus aus Bewertung, Diagnose, Rückmeldung und Förderung schließt sich an. Die gesamte Intervention bestand aus drei solcher Zyklen. Eine Woche nach der Intervention wurden das Fachwissen und die Einstellung zur Unterrichtsreihe und den

Feedback-Maßnahmen erhoben (Post). Vier Wochen später erfolgt der dritte Einsatz des Fachwissenstests (Follow-up).

#### *Auswertung*

Zur Überprüfung der ersten Forschungsfrage wurde der prozentuale Anteil richtiger Antworten im Fachwissenstest betrachtet. Daraus resultiert, dass die Lernenden aller Gruppen sowohl unmittelbar ( $n_{FB} = 60$ ,  $M_{pre} = .30$ ,  $M_{post} = .60$ ,  $p < .001$ ,  $\varphi = .83$ ;  $n_{LB} = 62$ ,  $M_{pre} = .30$ ,  $M_{post} = .59$ ,  $p < .001$ ,  $\varphi = .82$ ;  $n_{OR} = 63$ ,  $M_{pre} = .30$ ,  $M_{post} = .59$ ,  $p < .001$ ,  $\varphi = .85$ ) als auch nachhaltig ( $n_{FB} = 60$ ,  $M_{pre} = .30$ ,  $M_{Follow-up} = .59$ ,  $p < .001$ ,  $\varphi = .83$ ;  $n_{LB} = 62$ ,  $M_{pre} = .30$ ,  $M_{Follow-up} = .61$ ,  $p < .001$ ,  $\varphi = .85$ ;  $n_{OR} = 63$ ,  $M_{pre} = .30$ ,  $M_{Follow-up} = .60$ ,  $p < .001$ ,  $\varphi = .85$ ) signifikant dazulernen. Im Gruppenvergleich zeigt sich weder im unmittelbaren ( $p_{res} = .947$ ;  $\eta^2 = .001$ ) noch im nachhalten ( $p_{res} = .820$ ;  $\eta^2 = 0.002$ ) Fachwissenszuwachs ein Unterschied zwischen den drei Gruppen. Für die Bearbeitungsqualität konnten signifikante Unterschiede durch die konzipierte Unterrichtsreihe zwischen den drei Gruppen zugunsten der LB gegenüber der OR festgestellt werden bezogen auf das Gesamturteil der Schülerarbeitsmappe ( $n_{FB} = 60$ ,  $n_{LB} = 62$ ,  $n_{OR} = 63$ ;  $M_{FB} = 3.17$ ,  $M_{LB} = 3.21$ ,  $M_{OR} = 2.85$ ;  $p = 0.021$ ,  $\eta^2 = 0.042$ ). Hinsichtlich der Einstellung gegenüber der Unterrichtseinheit, Materialien und Feedback-Maßnahmen zeigt sich, dass die Lernenden aller Gruppen die Unterrichtseinheit und Materialien ähnlich positiv bewerten ( $n_{FB} = 60$ ,  $n_{LB} = 62$ ,  $n_{OR} = 63$ ;  $M_{FB} = 2.46$ ,  $M_{LB} = 2.66$ ,  $M_{OR} = 2.62$ ).

#### *Vier-Gruppen-Vergleich*

An einer der teilnehmenden Schule bestand die Möglichkeit neben den oben genannten drei Gruppen noch eine weitere vierte Gruppe (Vergleichsgruppe (VG)) während der Intervention zu generieren. Dies sind Lernende einer Parallelklasse, die von derselben Lehrkraft unterrichtet wurden. Die VG führt die Unterrichtsreihe wie die FB, LB und OR durch, jedoch ohne Kompetenztests zu schreiben, Feedbackmaterial zu erhalten bzw. den Lehrervortrag zu hören. Hierbei konnten signifikante Unterschiede zwischen den Gruppen FB und VG bzw. LB und VG festgestellt werden, und zwar jeweils zugunsten der FB- und LB-Gruppe bezogen auf die Pre-Post-Residuen ( $n_{FB} = 14$ ,  $n_{LB} = 14$ ,  $n_{OR} = 14$ ,  $n_{VG} = 16$ ;  $M_{FB} = 0.45$ ,  $M_{LB} = 0.26$ ,  $M_{OR} = -0.09$ ,  $M_{VG} = -0.79$ ;  $p = 0.008$ ,  $\eta^2 = 0.212$ ). Dieses Ergebnis ist mit Vorsicht zu betrachten, da nur eine kleine Stichprobe vorliegt. Hier deutet sich jedoch ein großer Effekt an, wobei der Unterschied zwischen den bisher eingesetzten Feedbackmaterialien zwar nicht stark zu sein scheint, wohl aber der Unterschied zu einer Gruppe ohne Feedback- bzw. Fördermaterial.

#### **Fazit und weitere Schritte**

Die ersten Ergebnisse der Hauptuntersuchung lassen darauf schließen, dass der große zeitliche Aufwand des individuellen, elaborierten Feedbacks wenig ertragreich erscheint gegenüber dem weniger aufwändigen Einsatz von Lösungsbögen. Aus dem Vergleich der vier Gruppen kann geschlossen werden, dass der zusammenfassende und wiederholende Lehrervortrag und die ausgeschärften Materialien, die alle Gruppen erhielten, möglicherweise Unterschiede im Fachwissen der Schülerinnen und Schüler der unterschiedlichen Gruppen ausgleichen konnten.

Im weiteren Verlauf werden die Ergebnisse zum Fachwissenszuwachs, der Einstellung und der Bearbeitungsqualität getrennt nach kognitiven Niveaus der Schülerinnen und Schüler ausgewertet. Zudem ist eine weitere Untersuchung geplant für das kommende Schuljahr 2017/2018, um die Kombination aus Feedback mit anschließendem Lehrervortrag zu untersuchen, wobei eine Realisierung der beiden Gruppen (Feedback mit Lehrervortrag und Feedback ohne Lehrervortrag) in jeder Klasse vorgesehen ist und der Umgang mit dem Feedbackmaterial durch Videoaufzeichnungen analysiert werden soll.

**Literatur**

- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21, 5-31.
- Cowie, B. & Bell, B. (2001). *The characteristics of formative assessment in science education*. Science Education. 85(5). S. 536-553.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Holländer, M. & Melle, I. (2012). Die Effektivität des Advance Organizers im Chemieunterricht der Sekundarstufe I. *Der mathematische und naturwissenschaftliche Unterricht*, 1, 44-52.
- Jürgens, E. & Lissmann, U. (2015). *Pädagogische Diagnostik*. Weinheim: Beltz Verlagsgruppe.
- Kultusministerkonferenz KMK (2002). PISA 2000 – Zentrale Handlungsfelder. Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern. Zugriff am 13.10.2017. Retrieved from <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2002/massnahmen.pdf>
- McMillan J. H. (2010). The Practical Implications of Educational Aims and Contexts for Formative Assessment. In: H. L. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 41-58). New York: Taylor & Francis.
- Ministerium für Schule und Weiterbildung des Landes NRW (2005). *Schulgesetz für das Land Nordrhein-Westfalen (SchulG-NRW)*. In der Fassung vom 15. Februar 2005 (GV. NRW. S. 102), zuletzt geändert am 13. November 2012 (GV. NRW. S. 514). Retrieved from <http://www.schulministerium.nrw.de/docs/Recht/Schulrecht/Schulgesetz/Schulgesetz.pdf>
- OECD (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Zugriff am 13.10.2017. Retrieved from <https://www.oecd.org/edu/cei/35661078.pdf>
- Rakoczy, K., Klieme, E., Bürgermeister, A. & Harks, B. (2008). The Interplay Between Student Evaluation and Instruction – Grading and Feedback in Mathematics Classrooms. *Journal of Psychology*, 216(2), 111-124.
- Rost, D. H., Sparfeldt, J. R. & Schilling, S. R. (2007). *DISK-GITTER mit SKSLF-8*. Differentielle Schulisches Selbstkonzept-Gitter mit Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten. Göttingen: Hogrefe.
- Sheppard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29 (7), 4-14.
- Stiggins, R. J. (2007). Conquering the Formative Assessment Frontier. In H. McMillan (Eds.), *Formative Classroom Assessment: Theory into practice* (pp. 8-28). New York: Taylor & Francis.
- Stiggins, R. J. & Chappuis, S. (2005). Putting Testing in Perspective: It's for Learning. *Principal Leadership*, 6(2), 16-20.
- Weiß, R. H. (1998). *Grundintelligenztest Skala 2 (CFT20)*. Göttingen: Hogrefe.
- Wiliam, D. & Leahy, S. (2007). A Theoretical Foundation for Formative Assessment. In H. McMillan (Eds.), *Formative Classroom Assessment: Theory into practice* (pp. 29-42). New York: Taylor & Francis.