

PROTOKOLL DES 3. TREFFENS DES NATIONAL STAKEHOLDER PANELS

Das dritte Treffen des deutschen National Stakeholder Panels (NSP) im Rahmen des ASSIST-ME-Projektes fand am 19.05.2016 im Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel statt.

Tagesordnung

13:00	Begrüßung und Projektstand
13:20	Diskussionsrunde I
13:55	Ergebnisse aus der Projektarbeit und einer Begleitstudie
14:35	Kaffeepause
15:00	Diskussionsrunde II
16:15	Ausblick
16:30	Ende des Treffens

Begrüßung und Projektstand

Nach der Begrüßung der Teilnehmerinnen und Teilnehmer stellt Olaf Köller zunächst noch einmal die Ziele und Fragestellungen des Projektes, die Projektphasen und -inhalte sowie das zentrale Konzept der formativen Diagnose und Bewertung vor.

Diskussionsrunde I: Diagnose und Bewertung im Unterricht und in der Aus- und Weiterbildung von Lehrkräften

In einer ersten Diskussionsrunde diskutieren die Panelmitglieder danach in zwei Gruppen die folgenden Fragestellungen: (1) Wie wird der Lernerfolg von Schülerinnen und Schülern in Ihrem Land erfasst? Bitte beschreiben Sie sowohl formative (z. B. Rückmeldungen der Lehrkräfte im täglichen Unterricht) als auch summative (z. B. Klassenarbeiten) Ansätze. Unterscheiden sich die Ansätze in Abhängigkeit von der Klassenstufe? und (2) Welche Rolle spielt das Wissen über formative und summative Diagnose und Bewertung in der Aus- und Weiterbildung von Lehrkräften?

Als Basis zur Diskussion der ersten Frage wird dem Panel folgende Tabelle zur Verfügung gestellt:

Summativ	Formativ
Int. Schulleistungsstudien <ul style="list-style-type: none"> • PISA (15-Jährige) • TIMSS (4) 	Diagnosebögen
Ländervergleiche (4 und 9)	Lernberichte, Lerntagebücher
Vergleichsarbeiten (VERA 3 und 8)	Wochenarbeitspläne
Klassenarbeiten	Portfolios
Zentrale Abschlussprüfungen	???
???	

Zunächst empfindet das Panel die Fragestellung an verschiedenen Punkten als zu ungenau. Die Fragestellung vermischt die Aspekte der Erfassung und Rückmeldung von Lernerfolg. Zudem ist nicht definiert, über welchen Zeitraum (einzelne Stunde, Woche, Monat, Jahr) und für wen (Schülerinnen und Schüler, Lehrkräfte, Eltern, System) Lernerfolg erfasst werden soll. Schließlich kann ‚Lernerfolg erfassen‘ sowohl im Sinne von einer Anbahnung oder Bewusstmachung nächster Schritte als auch im Sinne des Erreichens einer transparenten Bewertung und Beurteilung verstanden werden.

Bezüglich der Tabelle macht das Panel den Vorschlag, die Begriffe ‚summativ‘ und ‚formativ‘ durch ‚Leisten‘ und ‚Lernen‘ zu ersetzen. Die in der linken Spalte genannten Instrumente sind eindeutig Instrumente für Leistungssituationen (auf Individual-, Klassen- oder Systemebene). Die in der rechten Spalte genannten Instrumente sind eigentlich für Lernsituationen konzipiert – teilweise werden die dort gewonnenen Informationen allerdings summativ genutzt, wodurch Lern- und Leistungssituationen vermischt werden. Das Panel spricht sich eindeutig gegen eine solche Vermischung aus. Im Gegensatz zu Leistungssituationen zeichnen sich Lernsituationen dadurch aus, dass Fehler als potentielle Lernmöglichkeiten erwünscht sind; in Leistungssituationen werden Schülerinnen und Schüler jedoch versuchen, Fehler zu vermeiden. In Hinblick auf die genannten Instrumente würde das Panel in der rechten Spalte ‚Kompetenzraster‘ (auch als Möglichkeit zur Selbstreflektion) ergänzen. Außerdem sieht das Panel einen Übergangsbereich zwischen summativ und formativ. Hier wäre zum einen VERA einzuordnen (wobei die naturwissenschaftlichen Fächer in VERA nicht erfasst werden) sowie summative Tests, die durch formative Rückmeldungen ergänzt werden.

Prinzipiell empfindet das Panel den summativen Aspekt der Fragestellung als eher unproblematisch. Der formative Aspekt ist stark dadurch geprägt, dass formative Rückmeldung, insbesondere auf Individualebene, Zeit kostet und zeitlich eng an die Situation angebunden sein muss, damit sie für die Schülerinnen und Schüler effizient ist. Eine Möglichkeit, das Zeitproblem zu adressieren, besteht darin, Rückmeldung nicht auf Individualebene, sondern im Rahmen von Gruppenarbeit zu geben. Unterrichtsbeobachtungen zeigen, dass diese Form von Rückmeldungen, wenn auch nicht systematisch und standardisiert, im Unterricht vorkommt. Eine andere Möglichkeit könnte darin bestehen, Eigenlernzeiten im Schulalltag für diesen Zweck zu nutzen. Für das Problem der Individualisierbarkeit bieten sich computerbasierte Systeme an. Als Beispiele werden tutorielle, unterrichtsbegleitende Lernsysteme genannt, die auf Arbeitsschritte einzelner Personen reagieren, sowie im Unterricht eingesetzte Lernprogramme, die es der Lehrkraft ermöglichen, zeitnah, unkompliziert und fokussiert zu kontrollieren, wo ihre Schülerinnen und Schüler stehen (ohne beispielsweise dazu Materialien oder Hefte einsammeln zu müssen).

Als Basis der Diskussion der zweiten Frage dient ein Auszug aus den Standards für die Lehrerbildung: Bildungswissenschaften im Kompetenzbereich ‚Beurteilen‘ (KMK, 2014, S. 11-12). Mit Blick auf die Standards tauchte zunächst die Frage auf, ob in der ersten Phase der Lehramtsausbildung im Bereich Bewertung/Beurteilung ein zu großes Gewicht auf den Bildungswissenschaften liegt und damit die Fachspezifität vernachlässigt wird. Aus Sicht der Fachdidaktiken ist dieses jedoch nicht der Fall, sie sehen eine gute Passung zwischen bildungswissenschaftlichen und fachdidaktischen Inhalten. Als unglücklich wird hingegen die alleinige Verwendung des Begriffes ‚beurteilen‘ in den Standards bewertet, da der Begriff häufig mit Leistungsmessung assoziiert wird.

Aus der Sicht des Panels spielt der Bereich der Diagnose und Bewertung sowohl in der Aus- (1. und 2. Phase) als auch in der Weiterbildung eine zentrale Rolle. Eine wichtige – und in weiten Teilen offene – Frage ist jedoch, inwieweit die Umsetzung im Unterricht erfolgt und gelingt. Hierfür spielen verschiedene Faktoren eine Rolle. In Schleswig-Holstein z.B. gestaltet sich die Umsetzung auf Grund veralteter Lehrpläne, die den Lehrkräften gerade im prozessbezogenen Bereich keine vernünftige Orientierung bieten, schwierig. Es fehlen klare Kompetenzbeschreibungen, auf deren Basis Diagnose und Bewertung erfolgen können. Formative Diagnose und Bewertung erfordert Berufs- und Lebenserfahrung sowie ein solides fachliches Wissen. Die Lehrkraft muss in der Unterrichtssituation Probleme erkennen und die fachliche Sicherheit und Erfahrung besitzen, mit typischen Problemen umzugehen. Gerade jungen Lehrkräften, denen es an

Erfahrung und gegebenenfalls an Wissen mangelt, bietet die summative Diagnose und Bewertung eine Form von Sicherheit. In diesem Zusammenhang betont das Panel die Wichtigkeit der Akzeptanz der Verfahren durch die Lehrkräfte. Wird die Diskrepanz zwischen den theoretischen Inhalten, die den Lehramtsstudierenden in der ersten und zweiten Phase der Ausbildung vermittelt werden, und dem, was in der Praxis umzusetzen ist, zu groß, wird die Akzeptanz zwangsläufig sinken. In Berlin und Brandenburg ist die lernprozessbegleitende Diagnostik ein zentrales Thema in der Weiterbildung von Lehrkräften. Hier zeigt sich bei den Lehrkräften eine große Unsicherheit. Lehrkräfte benötigen nicht nur Instrumente zur Diagnostik, sondern sie müssen diese Instrumente und die dahinter liegenden Prinzipien auch verstehen. Eine zentrale Frage in diesem Zusammenhang ist beispielsweise, wie man die Anforderungen von Aufgaben variiert. Diese Entwicklungsarbeit sollte idealerweise von den Fachdidaktiken forschungsbasiert begleitet werden. Eine zweite Unsicherheit betrifft auch hier die Frage nach der tatsächlichen Umsetzung im Unterricht. Es reicht nicht, den Lehrkräften Materialien zur Verfügung zu stellen und diese in Fortbildungen zu besprechen, sondern die Umsetzung muss in der Praxis begleitet werden.

Das Panel betont hier nochmals den starken Einfluss der Schulen und der Fachgruppen in den Schulen, die Veränderungsprozesse unterstützen, aber auch behindern können. Aus- und Weiterbildung verfolgen langfristige Ziele, gleichzeitig müssen aber Schulentwicklungsprozesse in Gang gebracht werden. Eine Möglichkeit dazu stellen Kompetenzraster dar.

Ergebnisse aus der Projektarbeit und einer Begleitstudie

Im Rahmen von ASSIST-ME wurde in Deutschland das formative Diagnose- und Bewertungsverfahren der ‚Schriftlichen Rückmeldung‘ in den Fächern Chemie und Physik eingesetzt. Die Erprobungen fanden in drei aufeinanderfolgenden Erprobungsrunden in der Sekundarstufe I und II statt. In den Fallstudien zeigt sich, dass (a) die Lehrkräfte relevante Kompetenzaspekte in den Rückmeldungen ansprechen, (b) Begründungen der Einschätzungen jedoch häufig fehlen, insbesondere für schon erreichte Lernziele, (c) die Rückmeldungen überwiegend spezifische Hinweise auf nächste Lernschritte enthalten, (d) Schülerinnen und Schüler diese jedoch nur zum Teil nutzen. Als größte Stärke des Verfahrens erachteten die Lehrkräfte die Individualität, sowohl für die Schülerinnen und Schüler zur Kompetenzentwicklung als auch für die Lehrkräfte zur Unterrichtsentwicklung. Gleichzeitig bedeutet Individualität aber auch einen großen Zeit- und Arbeitsaufwand, der die Integration in den Unterrichtsalltag erschwert.

Befunde aus einer Begleitstudie, die die Lernwirksamkeit verschiedener Arten von schriftlichen Rückmeldungen untersucht, konnten zeigen, dass Rückmeldungen lernwirksam sind, wobei der größte Effekt für Rückmeldungen zu beobachten ist, die Informationen zu Lernzielen, aktuellem Lernstand und Verbesserungsmöglichkeiten enthalten. Gleichzeitig spielt aber auch der fachliche Inhalt und die Komplexität der Aufgabenstellung eine Rolle.

Zu den vorgestellten Ergebnissen gibt es im Panel verschiedene Fragen und Anregungen. Verschiedene Panelmitglieder bekunden Interesse an den im Projekt entwickelten Materialien, um sie in der Aus- und Weiterbildung von Lehrkräften einzusetzen. Die in Deutschland entwickelten Materialien können zur Verfügung gestellt werden; für die international entwickelten Einheiten und Instrumente wird das IPN den Wunsch an die internationale Projektleitung weitergeben. Großes Potential wird in dem in Deutschland nicht untersuchten Verfahren des Peer-Feedbacks gesehen. Die Erfahrungen der Projektpartner, die dieses Verfahren eingesetzt haben, zeigen jedoch, dass Schülerinnen und Schüler Zeit brauchen, dieses Verfahren zu erlernen. Zu der in der Begleitstudie beobachteten eingeschränkten Nachhaltigkeit der Lernwirksamkeit der Rückmeldungen gibt das Panel zu bedenken, dass langfristige Lernerfolge ein stetiges Lernangebot voraussetzen. Gleichzeitig regte das Panel an, das Follow up-Konzept zu überdenken. Das Vorhandensein eines nachhaltigen Effektes sollte dadurch definiert sein, inwieweit Schülerinnen und Schüler ihr zuvor erworbenes Wissen nach einem kurzen Input wieder reaktivieren können.

Ein Bereich, der im Projekt wenig adressiert, aber vom Panel als wichtig erachtet wird, ist der einer Kosten-Nutzen-Betrachtung. Rückmeldungen sind sehr zeitaufwändig, insofern muss die Lehrkraft eine Auswahl treffen. Fragen in diesem Zusammenhang wären z. B.: Wann und in welchem Bereich setze ich Rückmeldungen ein? Wie ausführlich gebe ich Rückmeldung? Wer profitiert von Rückmeldungen? In diesem Zusammenhang gibt das Panel den Rat, die Rückmeldungen und ihre Nutzung durch die Schülerinnen und Schüler nochmals detailliert zu analysieren, um Informationen zu gewinnen, welche Schülerinnen und Schüler von den Rückmeldungen profitieren, welche nicht und was mögliche Gründe hierfür sein könnten.

Diskussionsrunde II: Veränderungen der Diagnose- und Bewertungskultur

In der zweiten Diskussionsrunde geht es um die Frage, ob die Diagnose- und Bewertungskultur in Deutschland verändert werden sollte und, falls ja, wie solche Veränderungen realisiert werden können. Die Diskussion erfolgte erneut in zwei Gruppen. Die erste Gruppe fokussierte in ihrer Diskussion mehr auf die formative Diagnose- und Bewertung, während die zweite Gruppe sich schwerpunktmäßig mit

der Frage beschäftigte, ob Veränderungen in den Abschlussprüfungen notwendig sind, um die Kompetenzen zu erfassen, die SuS sowohl im Fach als auch fächerübergreifend erwerben sollen. Schließlich wurde das Panel gebeten, zu diskutieren, wo es eventuell Anknüpfungspunkte für weitergehende Untersuchungen sieht.

Nach Meinung des Panels ist es sowohl von der Schulart als auch von der Klassenstufe abhängig, ob und in welcher Form der Einsatz formativer Diagnose und Bewertung sinnvoll ist. Das Panel unterscheidet dabei zwischen defizit- und potentialorientierter Rückmeldung. Es wirft die Frage auf, ob die Tendenz eventuell zu stark Richtung potentialorientierter Rückmeldung geht. Damit könnte Schülerinnen und Schülern eine realistische Einschätzung ihrer Potentiale aber auch Defizite, wie sie beispielsweise für Übergänge in die berufliche oder universitäre Ausbildung wichtig ist, erschwert werden. Das Panel betont, dass das Erlernen der Fähigkeit, Rückmeldungen zu geben, Zeit erfordert – es ist daher wichtig, Lehrkräfte schrittweise daran heranzuführen. Ein möglicher Weg wäre, die Diagnosekompetenz zunächst im Studium in bereinigten, wenig komplexen Situationen (z. B. durch Unterrichtsvideos oder Wortprotokolle) zu fördern. In der zweiten Phase könnte die Auseinandersetzung mit Rückmeldungen im Sinne von Anregungen zum Weiterlernen dann systematisch intensiviert werden, woraus nach Einstieg in den Beruf erste Routinen, zunächst in bestimmten Bereichen entstehen. Damit sich daraus eine weitreichende Rückmeldekompetenz entwickelt, sind mehrere Jahre Unterrichtserfahrung vonnöten.

Eine Veränderung der Diagnose- und Bewertungskultur auf Systemebene erfordert allerdings zusätzlich eine Haltungsveränderung an den Schulen. Nach Meinung des Panels ist die Rückmeldekultur an Schulen in großen Teilen verloren gegangen – wenn es gelingt, sie wieder zu implementieren, wird sich das positiv auf Schulentwicklung, Unterricht und Lernerfolg auswirken. Auf das Gelingen solcher Prozesse hat die Schulleitung entscheidenden Einfluss. Gleichzeitig wird darauf hingewiesen, dass Schulen ein gewisses Maß an Ruhe brauchen, um sich solchen Fragen wie Rückmeldekultur zu widmen.

Auch an dieser Stelle spricht sich das Panel nochmal klar für eine Trennung von formativer und summativer Diagnose und Bewertung aus, um eine Vermischung von Lern- und Leistungssituationen zu vermeiden. Es betont allerdings auch, dass man auf die summative Diagnose und Bewertung nicht komplett verzichten kann. Für einen erfolgreichen Übergang in Ausbildung und Universität brauchen die Schülerinnen und Schüler eine Orientierung bezüglich ihrer Fähigkeiten und Defizite; hierzu sind regelmäßige, realistische Leistungsbeurteilungen nötig.

Dieses leitet zur zweiten Fragestellung, möglichen Veränderungen in den Abschlussprüfungen, über. Hier weist das Panel darauf hin, dass Veränderungen in den Abschlussprüfungen zwar einen steuernden Effekt haben, aber ohne einen vorgeschalteten Prozess, der diese Entwicklungen in die Schulen und den Unterricht integriert, bedeutet, dass mehrere Generationen von Schülerinnen und Schülern Prüfungen absolvieren müssen, auf die sie nicht angemessen vorbereitet wurden. Veränderungen in den Abschlussprüfungen muss daher immer ein Veränderungsprozess im Unterricht vorausgehen – in diesem Zusammenhang wird auf das Gutachten zur Vorbereitung des ersten SINUS-Programms (BLK, 1997) verwiesen. Die Veränderung der Prüfungsaufgaben sollte dann vorsichtig erfolgen, z.B. über Teilaufgaben, die die inhaltlichen Veränderungen der Bildungsstandards abdecken. Im Bereich der naturwissenschaftlichen Fächer, die im Rahmen von Abschlussprüfungen praktisch nur im Abitur vorkommen, bieten sich dabei z.B. Fragen zu experimentellen Methoden, zum Umgang mit Modellen oder zu naturwissenschaftlichen Denkweisen an. Eine Erfahrung aus Schleswig-Holstein unterstreicht die Wichtigkeit dieses Vorgehens. Vor einigen Jahren wurde das Mathematikabitur an beruflichen Schulen kompetenzorientiert ausgerichtet, ohne dass zuvor eine entsprechende Umorientierung im Kollegium und damit im Unterricht stattgefunden hatte. In der Folge waren die Schülerinnen und Schüler nicht in der Lage, die Aufgaben entsprechend zu lösen, woraufhin die Abiturnoten nachträglich korrigiert werden mussten. Bezüglich der Realisierung von Veränderungen rät das Panel, nicht allein politischen Einfluss zu nutzen, sondern diese sukzessive in inhaltliche Arbeitsprozesse zu integrieren.

Als mögliche Anknüpfungspunkte für weitergehende Untersuchungen nennt das Panel aus fachdidaktischer Sicht die Fortführung der Entwicklung praxistauglicher Rückmeldeinstrumente verbunden mit Studien zu deren Nutzung und Akzeptanz durch die Lehrkräfte. Nur, wenn man die Lehrkräfte erreicht und sie die Instrumente auch in der intendierten Form einsetzen, ist es sinnvoll, in einem nächsten Schritt auf Lernfortschritte der Schülerinnen und Schüler zu schauen.

Ein zweiter Anknüpfungspunkt betrifft eine stärkere Fokussierung auf die Kopplung von Diagnose und Förderung. In der Vergangenheit wurden häufig Diagnoseinstrumente in Schulen gegeben, ohne genau zu definieren, was damit gefördert werden soll. Eine Stärkung dieser Kopplung würde die Entwicklungsarbeit von Schulen und Unterricht maßgeblich unterstützen.

Ausblick

Das letzte Projektjahr in ASSIST-ME steht im Zeichen der Dissemination der Projektergebnisse. Sowohl auf nationaler als auch auf internationaler Ebene wurden Ergebnisse auf Tagungen vorgestellt und in Zeitschriften publiziert; in Schleswig-

Holstein fanden Arbeiten aus dem Projekt zudem in Fortbildungsveranstaltungen für Lehrkräfte Eingang. Für das nächste Jahr ist die Publikation eines Buches im Rahmen der ESERA Research Series des Springer-Verlags geplant. Außerdem soll im Herbst 2016 ein Treffen mit Stakeholdern aus den Projektländern in Brüssel stattfinden.

Zum Abschluss wird allen Mitgliedern des Panels wird für ihr Kommen und ihre Unterstützung über die gesamte Projektlaufzeit ausdrücklich gedankt. Die Rückmeldungen des Panels waren für die erfolgreiche Durchführung des Projektes sehr wertvoll.