

Joel Guttke

Diagnose von schulischem *underachievement* im Englischunterricht: Eine rekonstruktive Studie mit Englischlehrkräften der Sekundarstufe

Abstract

Diagnostische Kompetenz gilt als notwendige Voraussetzung für individuelle Förderung und adaptives Unterrichten. Dazu gehört auch die korrekte Einschätzung von Leistung und Leistungspotenzial, um intervenieren zu können, wenn Schüler*innen unterhalb ihres Leistungspotenzials verbleiben. Auf der Grundlage von episodischen Interviews und Gruppendiskussionen mit acht Englischlehrkräften der Sekundarstufe wird in diesem Beitrag mit Hilfe der dokumentarischen Methode rekonstruiert, welche impliziten Handlungslogiken die Teilnehmer*innen bei der Diagnose von Schulleistung leiten. Dabei wird erkennbar, dass (1) die Leistungsbewertung im Modus des Vergleichs innerhalb der Lerngruppe erfolgt, (2) schulisches *underachievement* als sichtbar oder unsichtbar konzeptualisiert wird und (3) erzwungene Kooperation als Interventionsmaßnahme Einsatz findet.

Diagnostic competence is regarded as an essential prerequisite for individual support and adaptive teaching. If teachers' diagnoses of academic achievement fail, their students are at risk of becoming academic underachievers. In this paper, data from episodic interviews and group discussions with eight secondary EFL teachers are analyzed from a reconstructive perspective. The analysis reveals that the participants (1) assess academic achievement in the mode of social comparison, (2) conceptualize academic underachievement as either visible or hidden, and (3) employ forced cooperation as an approach to address academic underachievement.

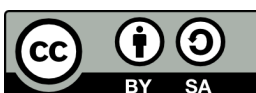
Schlagwörter:

underachievement, Englischunterricht, diagnostische Kompetenz, dokumentarische Methode, rekonstruktive Fremdsprachenforschung

underachievement, English as a foreign language, diagnostic competence, Documentary Method

I. Einleitung

Diagnostische Kompetenz ist für professionelles unterrichtliches Handeln notwendig, da sie Lehrkräften erst erlaubt, Unterricht zu gestalten, der den Bedürfnissen ihrer Schüler*innen entspricht. So definiert die Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) diagnostische Kompetenz unter der Überschrift „Diagnostik, Bewertung und Beratung“ (KMK, 2019: S. 4 f.) implizit als eines der elf Kernthemen der Lehrkräfteausbildung in Deutschland. Baumert und Kunter (2006) beschreiben diagnostische Kompetenz in ihrem Rahmenmodell professioneller Kompetenz als Aspekt fachdidaktischen Wissens. Nicht zuletzt ist die Passung zwischen Lernangelegenheiten und dem Lernstand der Schüler*innen im Englischunterricht ein zentrales



Moment in nahezu allen Ansätzen der Fremdsprachenerwerbsforschung (z. B. *comprehensible input*).

Wenn Schulleistung nicht oder nur unzureichend diagnostiziert wird und Leistungspotenzial unerkannt bleibt, besteht für Schüler*innen das Risiko des schulischen *underachievements* (UA). Betroffene Schüler*innen sind nicht dazu in der Lage, ihr Leistungspotenzial in der Schule vollständig auszuschöpfen; die Gründe dafür können vielfältig sein, wie beispielsweise ein Mangel an Motivation innerhalb oder fehlende Unterstützung außerhalb der Schule (vgl. Preckel & Baudson, 2013: S. 45). Im akademischen Diskurs erfolgt die Untersuchung von schulischem UA primär im Zusammenhang mit Hochbegabung. Montgomery (2020) kritisiert daran: „Most texts on underachievement [...] deal with it from the perspective of the needs of gifted children, those with high learning potential, when in fact children across the ability range are underachieving“ (S. 1). Ein solch weites Verständnis von schulischem UA¹ liegt auch diesem Beitrag zugrunde.

Nachfolgend wird UA sowie dessen Diagnose im Englischunterricht der Sekundarstufe fokussiert. Dazu werden die Ergebnisse einer rekonstruktiven Studie mit acht Englischlehrkräften der Sekundarstufe aus Nordrhein-Westfalen (NRW) berichtet. Sie veranschaulichen, wie der kollektive Habitus von Englischlehrkräften ihre diagnostische Praxis im Allgemeinen sowie die Identifikation von UA und die Implementation von Fördermaßnahmen diesbezüglich im Speziellen strukturiert. Verbunden mit dem Einsatz der dokumentarischen Methode ist ein Wechsel der Analyseperspektive weg von der Ebene des Inhalts und hin zu der Ebene der Performanz durch die Lehrkräfte. Somit untersucht der Beitrag nicht zwangsweise *was* die befragten Englischlehrkräfte diagnostizieren, sondern vielmehr *wie* sie Rückschlüsse über die Schulleistung sowie das Leistungspotenzial ihrer Schüler*innen ziehen und diese miteinander aushandeln.

Der Beitrag beginnt mit einer fachlichen Konkretisierung des Leistungsbegriffs durch die Beschreibung bildungspolitischer Vorgaben. Ein Einblick in Konzepte der Fremdsprachenerwerbsforschung ist notwendig, um die Relevanz diagnostischer Kompetenz von Englischlehrkräften für eine lernförderliche Unterrichtsgestaltung zu erläutern. Schließlich verdeutlicht der Übergang von der Primar- in die Sekundarstufe die fachspezifischen Herausforderungen, mit denen sich Englischlehrkräfte konfrontiert sehen. Es folgt die Darstellung empirischer Arbeiten über die diagnostische Kompetenz von Lehrkräften und UA. Abschließend werden basierend auf den berichteten Ergebnissen Implikationen für die Gestaltung des Englischunterrichts und der Englischlehrkräftebildung abgeleitet.

¹ UA kann grundsätzlich in allen Lebensbereichen auftreten. Wenn in diesem Artikel nachfolgend UA diskutiert wird, erfolgt die Diskussion stets im Bildungskontext, also mit explizitem Bezug auf schulisches UA.

2. Theoretischer Hintergrund zur Entwicklung der Forschungsfrage

2.1 Der Gemeinsame Europäische Referenzrahmen für Sprachen und die nationalen Bildungsstandards

Im Sinne der internationalen Vergleichbarkeit sind die Bildungsstandards für den fortgeführten Fremdsprachenunterricht in Anlehnung an den Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER) formuliert. Als bildungspolitisches Dokument stellt der GER „eine gemeinsame Basis dar für die Entwicklung von zielsprachlichen Lehrplänen, curricularen Richtlinien, Prüfungen, Lehrwerken usw. in ganz Europa“ (Europarat, 2001: S. 14). So attestieren die deutschen Schulabschlüsse, dass ihre Träger*innen interkulturell-kommunikative Kompetenzen erworben haben, die einem der gemeinsamen Referenzniveaus des GER zugeordnet werden können.

Innerhalb der Bildungsstandards folgt die KMK dem Weinertschen Kompetenzbegriff (vgl. Weinert, 2001: S. 27). Dieser wurde seit der Anwendung im Bildungsbereich kritisch diskutiert: Während beispielsweise Küster (2016) hervorhebt, der Kompetenzbegriff helfe dabei, die Individualität der Lerner*innen anzuerkennen, kritisiert Gunia (2012) das kompetenzorientierte Lehren und Lernen als weiteren Schritt hin zu einer Ökonomisierung von Bildung, die dem Kern eines humanistischen Bildungsideals widerspreche. Ungeachtet jeglicher Kritik bleibt festzuhalten, dass die Leistung von Schüler*innen – dem Kompetenzbegriff nach Weinert zufolge – anhand ihrer Performanz beurteilt wird. Diese Manifestation von Kompetenz durch ein ‚Zeigen können‘ verdeutlicht erneut das übergeordnete Anliegen, welches ursprünglich die Einführung der nationalen Bildungsstandards verursachte: die Ergebnisse unterrichtlichen Handelns – allgemeiner: Bildung – quantifizierbar zu machen.

Obwohl die Kompetenzbereiche in allen deutschen Kernlehrplänen ähnlich sind, werden sie in Folge des Bildungsföderalismus in den Bundesländern und an den einzelnen Schulformen mit unterschiedlichen pädagogischen Schwerpunkten umgesetzt (vgl. Elsner, 2018: S. 27–34). Ein Vergleich der Kernlehrpläne NRWs offenbart diese schulformspezifischen Konkretisierungen: Während beispielsweise an Hauptschulen ein Fokus auf dem mündlichen Gebrauch des Englischen liegt (vgl. MSWNRW, 2011: S. 8), werden an Gymnasien und Gesamtschulen sowohl Mündlichkeit als auch Schriftlichkeit fokussiert.

Eine solch unterschiedliche Gewichtung von Mündlichkeit und Schriftlichkeit ist insbesondere vor dem Hintergrund der Leistungserhebung und -beurteilung überraschend. Die Benotung von Schüler*innen setzt sich aus zwei Bereichen zusammen: Aus (1) schriftlichen Leistungsüberprüfungen und (2) der Beteiligung am Unterricht, die mündlich, schriftlich oder in Form von Projekten erfolgen kann. Eine der mindestens vier schriftlichen Leistungsüberprüfungen im Schuljahr darf – unabhängig von der Schulform – durch eine mündliche Leistungsüberprüfung ersetzt werden (vgl. MSB NRW, 2019: S. 41). Hinsichtlich der Leistungsbewertung werden Lehrkräfte dazu aufgefordert, neben der Performanz in der Prüfungssituation den Lernstand sowie die individuellen Stärken ihrer Schüler*innen zu berücksichtigen (vgl. MSB NRW, 2019: S. 40). Hieran wird deutlich, dass

die Umsetzung dieser komplexen Forderungen die diagnostische Kompetenz von Lehrkräften voraussetzt.

2.2 Konzepte der Fremdsprachenerwerbsforschung

Die Entwicklung sprachlicher und kognitiver Fähigkeiten bedingt sich gegenseitig. Der Fremdsprachenerwerb von Schüler*innen hänge von vielen internen und externen Faktoren ab, zu denen auch Interaktionen mit ihrem sozialen und institutionellen Umfeld zählen (vgl. Kersten, 2020). Befunde der (*Instructed*) *Second Language Acquisition research* tragen dazu bei, das Zusammenspiel dieser Faktoren zu verstehen und im institutionellen Rahmen durch die Formulierung pädagogischer Handlungsanweisungen zu verbessern (vgl. Loewen, 2015).

Eines der prominentesten Konzepte des Konstruktivismus ist die *zone of proximal development* (ZPD). Vygotsky (1978) definiert die ZPD als „the distance between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem-solving under adult guidance or in collaboration with more capable peers“ (S. 86). Schüler*innen werden nicht als passive Empfänger*innen von Lernaktivitäten betrachtet, sondern sie nehmen aktiv am Lernprozess teil. Indem Schüler*innen mit ihren Mitschüler*innen oder Lehrkräften interagieren, kann Lernen als Ko-Konstruktion von Wissen stattfinden. Im Englischunterricht findet das Konzept der ZPD aufbauend auf den Grundsteinen der Interaktion und des authentischen Sprachgebrauchs Ausdruck im Primat des *communicative language teaching*.

Als Teil seines *monitor model* argumentiert Krashen, es sei dem Fremdsprachenlernen zuträglich, wenn der *comprehensible language input*, mit dem sich Schüler*innen auseinandersetzen, das aktuelle Level ihrer *interlanguage* leicht übersteige ($i+1$; vgl. Lightbown & Spada, 2013: S. 105 f.). Obwohl diese Forderung hinsichtlich der im Englischunterricht zu erreichenden Lernprogression sinnvoll erscheint, wird sie vielfach kritisiert. Gass und Selinker (2008: S. 310) weisen darauf hin, es bleibe unklar, wie die aufeinanderfolgenden Fähigkeitslevel i und $i+1$ definiert seien und in welchem Ausmaß Schüler*innen mit dem *comprehensible language input* konfrontiert werden sollen. Um sicherzustellen, dass ihr *language input* für die Lerngruppe angemessen ist, können Englischlehrkräfte *scaffolding*, also „[...] forms of language support, learners are offered during the learning process“ (Kolb & Raith, 2018: S. 205), nutzen. Diese Form der Unterstützung kann sich beispielsweise sowohl auf die Unterrichtsstruktur und Formen der Leistungserhebung (*macro scaffolding*) als auch auf Unterrichtsmaterialien und Interaktionen im Klassenraum (*micro scaffolding*) beziehen.

Dies sind nur zwei von zahlreichen Beispielen, die verdeutlichen, dass eine Passung zwischen dem Sprach- und Unterrichtsangebot von Englischlehrkräften und den Bedürfnissen ihrer Schüler*innen notwendig ist. Eine solche Passung kann nur dann gelingen, wenn ausreichend Informationen über den Lernstand der Schüler*innen vorliegen und durch Englischlehrkräfte berücksichtigt werden. Dazu bedarf es diagnostischer Kompetenz.

2.3 Der Übergang von der Primar- zur Sekundarstufe

In Deutschland wurde der Fremdsprachenunterricht (gewöhnlich Englisch, selten Französisch) an Grundschulen im Jahr 2004 eingeführt. Seitdem beginnen Schüler*innen in 13 der 16 Bundesländer im dritten Schuljahr mit dem Englischunterricht. Seit 2021 lernen auch Grundschüler*innen in NRW, den Zielsetzungen des Masterplans Grundschule (MSBNRW, 2021) folgend, wieder ab dem dritten Schuljahr Englisch. Diese Unterscheidung zwischen Früh- und Spätstarter*innen wird politisch weiterhin kritisch diskutiert (vgl. Wilden & Porsch, 2020). Durch die Einführung des Englischunterrichts entstand an Grundschulen unter anderem ein Mangel an qualifizierten Englischlehrkräften. Bis heute wird ein Großteil der Grundschüler*innen von fachfremden Lehrkräften oder Lehrkräften mit Fortbildung unterrichtet (vgl. Porsch & Wilden, 2017: S. 65 f.).

Nach vier Jahren der Grundschule erfolgt für die Schüler*innen (mit Ausnahme von Berlin und Brandenburg) der Übergang zu einer Schule der Sekundarstufe. Im Vergleich zu anderen Ländern findet dieser Übergang in Deutschland frühzeitig statt. Mit einer Studie unter Englischlehrkräften in Baden-Württemberg zeigt Kolb (2011), dass der Übergang als eine Zeit der Diskontinuitäten wahrgenommen wird. Konkretisieren lassen sich diese Diskontinuitäten bspw. im Hinblick auf Kompetenzschwerpunkte bei der Unterrichtsgestaltung: Dominieren in Klasse 4 noch Arbeitsformen wie Lieder und Reime, Spiele sowie das Malen und Basteln, sind es in Klasse 5 hingegen die Vokabelarbeit und das Erklären grammatischer Strukturen (vgl. Kolb, 2011: S. 149 ff.). Dieser Herausforderung begegnet der Masterplan Grundschule (vgl. MSBNRW, 2021: S. 30) durch den zukünftigen Einbezug von Primar- und Sekundarlehrkräften in die Lehrplanarbeit. Des Weiteren haben zahlreiche Studien empirisch belegt, dass der frühe Übergang zu Bildungungerechtigkeit führt, von der Schüler*innen mit bestimmten Erstsprachen sowie bestimmter Ethnien und sozialer Klassen besonders betroffen sind (vgl. Valtin, 2020; van Ackeren, Klemm & Kühn, 2015: S. 78–90).

2.4 Diagnostische Kompetenz von (Englisch-)Lehrkräften

Obwohl es in der Fremdsprachenforschung bereits erste Versuche gab, diagnostische Kompetenz durch eine Fokussierung der kommunikativen Kompetenz von Schüler*innen fachspezifisch zu adaptieren (vgl. Kemmerer, 2019; Kubanek-German & Edelenbos, 2005), ist bisher kein Konsens erkennbar. Aus diesem Grund erfolgt hier ein überfachlicher Verweis auf diagnostische Kompetenz als

die Fähigkeit, die im Lehrberuf anfallenden diagnostischen Aufgabenstellungen erfolgreich zu bewältigen, und [...] die Qualität der dabei erbrachten Diagnoseleistungen. Ziel ist es, Informationen über Lernergebnisse, Lernvoraussetzungen und Lernvorgänge von Schülerinnen und Schülern zu gewinnen, die für verschiedene pädagogische Entscheidungen [...] genutzt werden können (Schrader, 2013: S. 154).

Die Tatsache, dass diagnostische Kompetenz primär über ihre Ziele definiert wird, bekräftigt die Unschärfe des Konzepts. Diagnostische Methoden, als ein beobachtbares Resultat diagnostischer Kompetenz, können auf einem Kontinuum von informeller zu formeller Diagnostik platziert werden: Während informelle Diagnosen weniger zeitintensiv und

einfacher in den Unterrichtsalltag zu integrieren sind, zeichnen sich formale Diagnosen in der Regel durch ein höheres Maß an Güte aus (vgl. Schrader, 2013: S. 155, 161). Außerdem weisen empirische Studien zu *teacher expertise* (vgl. Berliner, 1994: S. 152 ff.) darauf hin, dass diagnostische Kompetenz – als eine Facette von professioneller Kompetenz – domänenspezifisch ist.

Ein hoher Grad an diagnostischer Kompetenz verbessert die Genauigkeit der Erwartungen, die Lehrkräfte aufgrund vorliegender Informationen an die Schulleistung ihrer Schüler*innen stellen. Solche Leistungserwartungen sind insbesondere im Kontext informeller Diagnostik fehleranfällig, da sie von den Eigenschaften der Schüler*innen (z. B. das durchschnittliche Leistungsniveau einer Lerngruppe) und Lehrkräfte (z. B. ein Fokus auf leistungsstarke Schüler*innen) beeinflusst werden.

Um die Genauigkeit der diagnostischen Kompetenz von Lehrkräften empirisch zu beurteilen, werden die Bewertungen der Schulleistung von Schüler*innen gewöhnlich mit den Ergebnissen der Schüler*innen in einem standardisierten Leistungstest verglichen. Dadurch ist es möglich, drei Indikatoren diagnostischer Kompetenz zu berechnen: die Niveauelemente, die Differenzierungskomponente und die Rangkomponente (vgl. Schrader & Helmke, 1987: S. 30 f.). Zahlreiche Studien belegen, dass Lehrkräfte dazu tendieren, die durchschnittliche Leistung ihrer Schüler*innen in einer Lerngruppe zu überschätzen (vgl. Spinath, 2005; Südkamp, Möller & Pohlmann, 2008).

In einer Meta-Analyse fassen Südkamp, Kaiser und Möller (2012) die empirischen Befunde zu der Genauigkeit von Lehrkräfteurteilen auf Grundlage von 75 Studien aus den Jahren 1989–2009 zusammen. Korrelationsanalysen zeigen einen positiven Zusammenhang ($r = .63$) zwischen Lehrkräfteurteilen und den Schulleistungen von Schüler*innen. In der Gesamtschau werden einerseits interpersonelle Unterschiede zwischen den untersuchten Lehrkräften deutlich; andererseits sind durchschnittlich 60% der Varianz in Lehrkräfteurteilen nicht auf die Leistung der Schüler*innen zurückzuführen. Budson, Fischbach und Preckel (2016: S. 155) klären diesen Varianzanteil auf, indem sie zeigen, dass das Bildungsniveau der Eltern sowie der Grad der Bekanntheit zwischen Lehrkräften und Schüler*innen stark mit den Lehrkräfteurteilen korrelieren. Außerdem besteht eine positive Korrelation ($r = .58$) zwischen den Lehrkräfteurteilen über die kognitiven Fähigkeiten ihrer Schüler*innen und deren Intelligenz. Diese Ergebnisse bestärken die Beobachtung, dass Lehrkräfte bei der Bewertung der Leistungen ihrer Schüler*innen auf ihr Wissen über deren kognitive Fähigkeiten zurückgreifen (vgl. Rost & Hanses, 1997). Demzufolge können Lehrkräfte die kognitiven Fähigkeiten ihrer Schüler*innen nur so weit einschätzen, wie sie in ihrer Performanz Ausdruck finden (vgl. Schrader & Helmke, 2001).

Lamprecht (2007) rekonstruiert auf Grundlage von Gruppendiskussionen mit zwölf Lehrkräften der Sekundarstufe den Habitus, welcher die Leistungsbeurteilung als eine Komponente ihrer diagnostischen Kompetenz strukturierte. Einer der zentralen Befunde besteht darin, dass die Orientierungsrahmen der Lehrkräfte in Abhängigkeit von dem sozialen Milieu ihrer Schüler*innen variieren. Lehrkräfte, die an Schulen in sozial deprivierter Lage arbeiten, vermeiden jegliche Form von Leistungsbewertung. Dies dokumentiert sich in der Stigmatisierung von Schüler*innen und einem geringen Selbstwirksamkeits-

konzept der Lehrkräfte. Lehrkräfte, die Schüler*innen der Mittelschicht unterrichten, ermutigen ihre Schüler*innen hingegen explizit zur Leistungsentwicklung (S. 148).

Hinsichtlich der domänenspezifischen diagnostischen Kompetenz von Englischlehrkräften ist die Befundlage begrenzt. Leucht, Tiffin-Richards, Vock, Pant und Köller (2012) untersuchen die diagnostische Kompetenz von Lehrkräften in Deutschland hinsichtlich des Leseverstehens in der Fremdsprache Englisch. Dabei waren die Autor*innen besonders an der Effektivität des GER als diagnostisches Tool für eine objektive(re) Leistungsbewertung interessiert. 16% der 56 teilnehmenden Englischlehrkräfte gaben an, sehr vertraut mit dem GER zu sein während 12,5% angaben, den GER bereits in der Vergangenheit zu diagnostischen Zwecken verwendet zu haben. Im Gegensatz dazu waren 68% nur vage mit dem GER vertraut und 16% der Lehrkräfte gaben an, noch nie vom GER gehört zu haben (S. 171). Diese Bestandsaufnahme ist allein vor dem Hintergrund der in Abschnitt 2.1.1 geschilderten Relevanz bildungspolitischer Dokumente alarmierend. Im Durchschnitt unterschätzten die Lehrkräfte mit ihrem Urteil die Leseverstehenskompetenz der Schüler*innen in allen 56 Klassen. Die durchschnittliche Differenzierungskomponente deutet jedoch daraufhin, dass die Lehrkräfte die Varianz in der Leseverstehenskompetenz überschätzten (S. 172). Zudem verfügten Englischlehrkräfte, die den GER in der Vergangenheit bereits im Rahmen der Leistungsbewertung eingesetzt hatten, über eine qualitativ höhere diagnostische Kompetenz (S. 173).

Die unerwartet niedrigen durchschnittlichen Lehrkräfteurteile in der Studie von Leucht et al. (2012) implizieren, dass die teilnehmenden Englischlehrkräfte zu wenig von ihren Schüler*innen erwarten (S. 175). Die hier vorliegende Diskrepanz zwischen dem Leistungsniveau der Schüler*innen und den Lehrkräfteurteilen kann verhindern, dass Lehrkräfte innerhalb der ZPD ihrer Schüler*innen unterrichten oder angemessenes *scaffolding* bereitstellen. Eine solche Unterschätzung kann dazu führen, dass sich Schüler*innen unterfordert fühlen und weniger leisten als sie könnten.

2.5 Schulisches (*under-*)*achievement* im Englischunterricht

Schulleistung ist in einem komplexen Modell von gegenseitigen Abhängigkeiten verortet, die Helmke (2015) in seinem Angebot-Nutzungs-Modell schulischen Lehrens und Lernens zusammenfasst. Diesem Modell liegt die Annahme zugrunde, die Effekte von Lerngelegenheiten, die Lehrkräfte anbieten, seien multidimensional und hingen in hohem Maße davon ab, ob Schüler*innen diese Gelegenheiten nutzen. Das Angebot-Nutzungs-Modell ist vereinbar mit der Definition schulischer Leistung nach Ingenkamp und Lissmann (2008): „Unter Schulleistung versteht man zusammengefasst die von der Schule initiierten Lernprozesse und Lernergebnisse der Schüler“ (S. 131). Diese Definition sieht die Performanz von Schüler*innen als Basis von Schulleistung ohne dabei die Rolle der Lehrkräfte in der Mitwirkung an dieser Performanz sowie die Beziehung zwischen Lehrkräften und Schüler*innen zu vernachlässigen.

Als ein Aspekt der Wirkungen im Angebot-Nutzungs-Modell manifestiert sich Schulleistung in den Kompetenzen der Schüler*innen in Form deklarativen, prozeduralen und metakognitiven Wissens (Weinert, 2001). Intelligenz korreliert moderat ($r = .50$) mit

Schulleistung. Zusammen mit Vorwissen ist Intelligenz der stärkste Prädiktor von Schulleistung (vgl. Brühwiler, Helmke & Schrader, 2017). Die motivationalen und volitionalen Facetten von Schulleistung umfassen hingegen die Fähigkeit, sich selbst zu motivieren sowie Metakognition, die als Voraussetzung für die Anwendung von Lernstrategien gilt.

McCoach und Siegle definieren UA als Diskrepanz zwischen Potenzial und Performanz (vgl. Siegle, 2018: S. 286). Um von UA zu sprechen, muss eine solche Diskrepanz stabil über einen bestimmten Zeitraum bestehen und darf nicht aus situativen, temporären Umständen resultieren. Als Indikator für das Potenzial von Schüler*innen wird gewöhnlich der Intelligenzquotient (IQ) herangezogen, der mithilfe eines standardisierten Intelligenztests gemessen wird (vgl. Sparfeldt & Rost, 2012: S. 436). Vor dem Hintergrund des zuvor beschriebenen Einflusses von Intelligenz auf Schulleistung und der hohen psychometrischen Qualität von Intelligenztests erscheint dieses Vorgehen gerechtfertigt. Hinsichtlich der Bewertung der Performanz wird in der Bildungsforschung in einem Großteil der Studien auf Schulnoten oder Punkte in standardisierten Schulleistungstests verwiesen.

Unabhängig von der Art und Weise, wie Potenzial und Performanz gemessen werden, unterscheidet man unterschiedliche Ansätze zur Definition der Diskrepanz zwischen beiden Variablen. Konservative Ansätze spezifizieren *cut-off* Werte, denen zufolge *underachieving* Schüler*innen ein Potenzial in Form eines IQs oberhalb eines bestimmten Perzentils und eine Performanz unterhalb eines bestimmten Perzentils erzielen. Dieses Diskrepanzkriterium wurde durch die Verwendung von Regressionsmodellen geöffnet, indem dann von Diskrepanz gesprochen wird, wenn das Performanz-Maß mindestens eine Standardabweichung unterhalb des Wertes liegt, der sich prädiktiv aus dem IQ der Schüler*innen ergibt. In Deutschland wurden nur wenige Studien zu UA durchgeführt, die versuchen, das Diskrepanzkriterium an das mehrgliedrige Schulsystem anzupassen (vgl. Uhlig, Solga & Schupp, 2009). Eine Einschränkung dieses Ansatzes besteht darin, dass sich die Autor*innen auf die drei traditionellen Schulformen (Hauptschule, Realschule und Gymnasium) beschränken und damit hybride Formen wie Gesamt- oder Sekundarschulen ausschließen.

Die Vagheit des Konstrukts, die mit den vielfältigen Definitionen von UA verbunden ist, hat wiederholt zu kritischen Rezeptionen geführt (vgl. Stamm, 2011). Im Rahmen des vorliegenden Beitrags wird ein stärker qualitatives Verständnis von UA zugrunde gelegt. Informationen hinsichtlich des Potenzials und der Performanz von Schüler*innen werden Interviews und Gruppendiskussionen mit Englischlehrkräften der Sekundarstufe entnommen. Da das Forschungsinteresse auf dem impliziten Wissen liegt, auf das Lehrkräfte bei der Diagnose von UA zurückgreifen, müssen diese Diagnosen nicht im zuvor beschriebenen, strengen Sinne korrekt sein. In der Tat ist auch die Diagnose von Schüler*innen, die fälschlicherweise als *underachievers* eingestuft wurden, ertragreich für das Erkenntnisinteresse. Preckel und Baudson (2013: S. 38 f.) heben hervor, dass Lehrkräfte je nach Berufserfahrung Zugriff auf eine große Referenzgruppe an Schüler*innen haben, welche die Genauigkeit ihrer Urteile begünstigen kann. Dieses Verständnis von UA erlaubt die Berücksichtigung von Schüler*innen entlang des gesamten Schulleistungskontinuums,

denn „particular types of learning disability, such as in language, [...] may lower IQ subtest scores just as a poor socio-cultural environment can“ (Montgomery, 2020: S. 7).

3. Empirische Studie zur Diagnose von schulischem *underachievement* im Englischunterricht der Sekundarstufe

3.1 Forschungsfrage

Zunehmend heterogene und zieldifferent unterrichtete Lerngruppen erfordern einen diagnostischen Fokus auf individuelle Leistungsentwicklungen und entkräften somit den Vergleich innerhalb der Lerngruppe als Mittel zur Leistungsbewertung. Zudem forciert das mehrgliedrige Schulsystem eine Förderung basierend auf der Illusion leistungshomogener Lerngruppen. Bei UA scheinen Strategien zur individuellen Förderung wie der ZPD und *scaffolding* zu scheitern, da Schüler*innen trotz entsprechender kognitiver Fähigkeiten erwartungswidrig niedrige Leistungen erzielen. Eine Differenzierung zwischen Potenzial und Performanz bleibt aus; es wird zur Selbstverständlichkeit, Schüler*innen seien dazu in der Lage, ihr Potenzial in Leistung umzusetzen. Während Performanz immer auf einem gewissen Maß an Potenzial basiert, wird die Umsetzung von Leistungspotenzial in Performanz durch verschiedene Faktoren vermittelt – eine Annahme, die sich auch im Angebot-Nutzungs-Modell widerspiegelt. Der vorliegende Beitrag trägt zu der Beantwortung der folgenden Forschungsfrage bei: Welcher habitualisierten Handlungslogik folgt das Verhalten von Englischlehrkräften der Sekundarstufe bei der Diagnose von UA, also im Falle unerwartet niedriger Schulleistung?

3.2 Methodologische Grundannahmen

Die dokumentarische Methode ging als forschungspraktisches Verfahren aus den sensussoziologischen Beiträgen Mannheims und Bourdieus hervor. Sie ermögliche die analytische Trennung generalisierten kommunikativen Wissens sowie handlungsleitenden konjunktiven Wissens (vgl. Bohnsack, 2014). Kommunikatives Wissen umfasst Wissensanteile, die für Individuen reflexiv zugänglich sind. Wenn Englischlehrkräfte zu ihrer Unterrichtsplanung befragt werden, sind sie gewöhnlich dazu in der Lage, ihre Planungsentscheidungen unter Rückgriff auf ihr Wissen über Prinzipien des Fremdspracherwerbs zu begründen. Im Gegensatz dazu sind konjunktive Wissensanteile, die aus der Teilhabe an bestimmten *communities of practice* resultieren und tägliche Entscheidungen und Handlungen strukturieren, nicht reflexiv zugänglich. Anteile konjunktiven Wissens seien exklusiv und blieben Individuen außerhalb einer bestimmten *community of practice* unzugänglich: „Das Wissen, das notwendig ist, um ‚dazuzugehören‘, ist in die Handlungspraxis eingelassen, besteht mithin im ‚Mitmachen-Können‘“ (Przyborski & Wohlrab-Sahr, 2014: S. 286).

Das Anliegen der dokumentarischen Methode bestehe darin, Zugang zu diesen konjunktiven Erfahrungsräumen zu erhalten, indem Texte, Interaktionen und andere Medien analysiert werden, die aus ihnen heraus entstehen (vgl. Bohnsack, 2014). Die steigende

Anzahl an Forschungsprojekten kann als Indikator dafür gesehen werden, dass das Erkenntnispotenzial rekonstruktiver Forschung (vgl. Bonnet, 2020) auch in der Fremdsprachenforschung zunehmend Anerkennung findet.

Um die für die Beantwortung der Forschungsfrage relevanten Aspekte impliziten Wissens rekonstruieren zu können, war es notwendig Daten zu sammeln, in denen Englischlehrkräfte der Sekundarstufe ihre alltäglichen pädagogischen Routinen hinsichtlich der Diagnose von Schulleistung offenbaren. Diese Erzählungen enthalten Spuren des impliziten Wissens der Lehrkräfte, die sich in der Interaktion mit ihren Kolleg*innen und Schüler*innen als Teil ihres professionellen Handelns äußern: „Die Gruppe ist somit nicht der soziale Ort der Genese und Emergenz, sondern derjenige der Artikulation und Repräsentation [...] kollektiver Entscheidungen“ (Bohnsack, 2000: S. 378).

3.3 Datenerhebung mit Hilfe episodischer Interviews und Gruppendiskussionen

Die Rekrutierung von Teilnehmer*innen begann im August 2020. Um das deutsche Bildungssystem in seiner Vielfalt abzubilden, wurde zunächst versucht, Englischlehrkräfte von allen Schulformen der Sekundarstufe (Haupt-, Real-, Gesamt- und Sekundarschulen sowie Gymnasien) miteinzubeziehen. Dazu wurden randomisiert 200 Schulen der Sekundarstufe (jeweils 40 Schulen pro Schulform) in NRW kontaktiert. Fünf Lehrkräfte von einer Realschule und einem Gymnasium erklärten sich zu der Teilnahme bereit. In einem zweiten Schritt wurden informell weitere Lehrkräfte angefragt, was in einer selektiven Auswahl von vier Fällen (A–D) mit insgesamt acht Teilnehmer*innen (Tab. 1) resultierte. Aufgrund der COVID-19-Pandemie wurde den Teilnehmer*innen angeboten, die Datenerhebung alternativ online in Form einer Videokonferenz durchzuführen; dieses Angebot akzeptierten die Lehrkräfte aus Fall B und Fall D.

Fall	Schulform	Teilnehmer*innen (selbst gewählte Pseudonyme)	Länge der Gruppendiskussion/des Interviews
A	Realschule	L1	00:55:00
B	Gesamtschule	L2	00:25:00
C	Gymnasium	L3, L4, L5, L6	00:40:00
D	Sekundarschule	L7, L8	01:00:00

Tab. 1: Übersicht der teilnehmenden Englischlehrkräfte

In Vorbereitung auf die Gruppendiskussionen und Interviews wurde die Forschungsfrage – den Ausführungen von Helfferich (2011) folgend – in einen Leitfaden überführt. Dieser umfasst vier thematische Abschnitte: (1) diagnostische Praxis und Methoden der Englischlehrkräfte, (2) schulische Leistungsindikatoren, (3) schulisches UA und (4) Heterogenität der Lerngruppen. Vorformulierte Redeaufforderungen halfen dabei, Fragen mit starken Präsuppositionen zu vermeiden. Nahezu alle Redeaufforderungen des Leitfadens zielen darauf ab, konkrete Erzählungen über die Erfahrungen der Teilnehmer*innen zu initiieren (z. B. Redeaufforderung zu Abschnitt 1 des Leitfadens: „Erzählen Sie mir bitte –

gerne anhand eines konkreten Beispiels – wie Sie den Lernstand Ihrer Schüler*innen im Englischunterricht feststellen.“). Dieser Ansatz basiert auf Schützes Annahme, dass Erzählungen als kognitiver Repräsentationsmodus im Gegensatz zu Beschreibungen oder Argumentationen den höchsten Grad an Übereinstimmung mit den ursprünglichen Erfahrungen der Erzähler*innen zeigen. Przyborski und Wohlrab-Sahr (2014: S. 80 f.) begründen diese Annahme mit einem Verweis auf die Zugzwänge des Erzählens.

Da es – anders als ursprünglich geplant – nicht möglich war Teilnehmer*innen für fünf Gruppendiskussionen zu finden, wurden mit zwei Lehrkräften (L1 und L2), die sich als einzige aus ihren Kollegien zur Teilnahme bereit erklärten, episodische Interviews geführt. Diese Form des Interviews kombiniert Elemente leitfadengestützter und erzählgenerierender Interviews, um das semantische und episodische Wissen der Interviewten zu erheben (vgl. Flick, 2011). Obwohl sich die resultierenden Daten hinsichtlich der Diskursorganisation fundamental von denen der Gruppendiskussionen unterscheiden, beinhalten sie äquivalente Sequenzen und wurden somit als zusätzliche Datenquellen in die Analyse einbezogen.

3.4 Datenauswertung mit Hilfe der dokumentarischen Methode

Die Rekonstruktion des impliziten Wissens der Englischlehrkräfte folgte den Vorschlägen von Bohnsack und Pfaff (2010) und Nohl (2013). Erstens wurde das erhobene Datenmaterial gesichtet, um relevante Abschnitte für die Transkription zu identifizieren. Dabei wurden Abschnitte als relevant betrachtet, wenn sie (1) Themen enthielten, von denen bereits zuvor in empirischer Forschung berichtet wurde, (2) aus einer hohen Interaktionsdichte zwischen den Teilnehmer*innen resultierten oder (3) wiederholt im Laufe des Gesprächs auftraten (vgl. Nohl, 2013: S. 40). Die Transkription erfolgte mit der Software MAXQDA (VERBI Software, 2019) und dem Transkriptionssystem *Talk in Qualitative Social Research* (vgl. Bohnsack, 1989: S. 387 f.). Sie ermögliche den Zugang zu der sozialen Realität der Teilnehmer*innen ohne das Gesprochene bereits durch die Interpretation der Wissenschaftler*innen zu filtern (vgl. Bohnsack & Pfaff, 2010: S. 16).

Zweitens erfolgte die formulierende Feininterpretation. Dabei wurden die Transkripte den Gesprächsthemen der Englischlehrkräfte entsprechend unterteilt und paraphrasiert. Der Fokus lag darauf, *was* die Englischlehrkräfte in Reaktion auf die Redeaufforderung zur Sprache brachten (immanenter Sinngehalt). In diesem Interpretationsschritt kristallisierten sich bereits Gesprächsthemen heraus, die in allen vier Fällen in ähnlicher Weise diskutiert wurden (z. B. Ausrichtung des Unterrichts am durchschnittlichen Leistungsniveau der Lerngruppe).

Drittens wurde durch die reflektierende Feininterpretation herausgearbeitet, *wie* die Englischlehrkräfte die Gesprächsthemen miteinander verhandelten (dokumentarischer Sinngehalt). Bei der Untersuchung der Diskursorganisation wurden vier Textsorten voneinander unterschieden: Erzählungen, Beschreibungen, Argumente und Bewertungen. Diese Textsortentrennung ist notwendig, da in Erzählungen und Beschreibungen Wiedergegebenes – den obigen Ausführungen folgend – mit dem ursprünglich Erlebten der Englischlehrkräfte am stärksten übereinstimmt. Argumentationen und Bewertungen

hingegen zeichnen sich durch einen hohen Grad an Reflexion aus und korrespondieren folglich stark mit dem kommunikativen Wissen der Sprecher*innen. Weiterhin wurde berücksichtigt, wie die Englischlehrkräfte Diskurse zu einzelnen Gesprächsthemen eröffnen, strukturieren und schließen. Diesbezüglich galt die Aufmerksamkeit insbesondere positiven und negativen Vergleichshorizonten, mit deren Benennung sich die Englischlehrkräfte positionierten.

Diese Arbeitsschritte trugen zu der Rekonstruktion der handlungsleitenden Orientierungen der befragten Englischlehrkräfte bei. Die vermuteten impliziten Regelmäßigkeiten, die sich in den Orientierungen der Lehrkräfte dokumentierten, wurden durch fallimmanente und fallübergreifende, vergleichende Sequenzanalysen überprüft. So wurde „[...] der Standort der Forscherin [...] dann nur noch ein Dreh- und Angelpunkt unter anderen [ist]“ (Przyborski & Wohlrab-Sahr, 2014: S. 282). Zusätzlich diskutierte der Autor seine Interpretationen mit zwei Kommilitoninnen zu drei Zeitpunkten während des Auswertungsprozesses.

4. Ergebnisse

Aufgrund des Beitragsumfangs beschränkt sich die Ergebnisdarstellung auf zentrale Ausschnitte des Datenmaterials und verletzt damit die Forderung nach einer „Historisierung“ des Deutungsprozesses“ (Reichert, 1992: S. 346). Wenn nachfolgend auf die Fälle A–D verwiesen wird, ist damit jeweils eine der vier Erhebungssituationen (s. Tab. 1) gemeint. Durch den Vergleich der Charakteristika aller vier Fälle im Rahmen der komparativen Sequenzanalyse konnten drei Orientierungen (I–III) herausgearbeitet werden, die relevant für die Diagnose unerwartet niedriger Performanz sind. Abbildung 1 illustriert diese Orientierungen, deren Genese in den drei folgenden Kapiteln detailliert beschrieben wird, anhand von repräsentativen Zitaten aus dem Datenmaterial. Die drei Orientierungen beschreiben die pädagogische Praxis der Englischlehrkräfte hinsichtlich der Leistungsdiagnose im Allgemeinen (Kap. 4.1), der Identifikation von UA (Kap. 4.2) und der Intervention im Falle von UA (Kap. 4.3).

I. Beurteilung schulischer Leistung im Modus des Vergleichs innerhalb der Lerngruppe



„also ich diagnostizier da schon eher über die Dinge die besonders gut laufen oder wo ich sag wir packen da immer wieder ins Leere (.) die Mitte eigentlich nicht so sehr“

(Fall C, Teiltranskript I)

II. Konzeptualisierungen von *underachievement* als (un-)sichtbar



„n Schüler der grundsätzlich großes Interesse an der Fremdsprache::: gezeigt hat und auch (.) großes Potenzial aba insgesamt verhaltensauffällig eh: war“

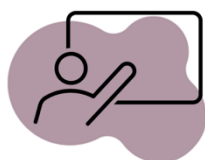
(Fall B, Teiltranskript II)



„d-der Bildungshintergrund der Eltern is jetzt eh::: auch nicht eh::: so dass man das als ideale Voraussetzung eh: erstmal sehen würde und dieses Mädchen is aber richtig richtig clever“

(Fall C, Teiltranskript II)

III. Intervention bei *underachievement* im Modus der erzwungenen Kooperation



„hat ganz massiv dann auch geschwänzt (.) und eh::: ich hab hinterher telefoniert was das Zeug hält [...] und er hat tatsächlich auch von mir nicht den Abschluss bekommen den er verdient hätte“

(Fall A, Teiltranskript III)

Abb. 1: Ergebnisse der sinngenetischen Typenbildung (eigene Darstellung)

4.1 Beurteilung schulischer Leistung im Modus des Vergleichs innerhalb der Lerngruppe

Um nachzuvollziehen, auf welches implizite Wissen die Englischlehrkräfte im Falle unerwartet niedriger Performanz zurückgreifen, ist es notwendig zu verstehen, wie sie Schulleistung im Allgemeinen diagnostizieren. Die anfänglichen Redeaufforderungen zielten daher auf Erzählungen der Lehrkräfte über ihr unterrichtliches Handeln zur Lernstandsdiagnose ab. In allen vier Fällen führte dies zu der Diskussion einer Schüler*innengruppe, die in Fall C als „die Mitte“ bezeichnet wird:

- 70 L4: [...] eh::: oder was können sie eben besonders gut also ich diagnostizier da
71 schon eher über die Dinge die besonders gut laufen oder wo ich sag wir packen
72 da immer wieder ins Leere (.) die Mitte eigentlich nicht so sehr
73 L3: L Mh::: ((zustimmend)) richtig sehe
74 ich auch so die Mitte is immer halt so eher dünn besetzt
75 L4: L Das is so diese ok Mitte @(.).@
76 L6: Ja das sind ja auch die um die man sich so am wenigsten eigentlich kümmern
77 muss ne? das sind so die die **so klakomm** ne? das sind nicht die den man viel
78 helfen muss das sind nicht die die als erstes fertig sind und denen langweilig wird
79 ne? das sind so die (.) auf den man so ja (.) auf die das am besten immer so drauf
80 passt

L4 definiert „die Mitte“ implizit als Restklasse, indem sie besonders leistungsschwache und -starke Schüler*innen fokussiert (Z. 70–72). Zudem erfolgt durch die Bezeichnung eine Homogenisierung der Schüler*innen aufgrund ihrer durchschnittlichen Schulleistung, wobei individuelle Charakteristika in den Hintergrund treten. Dieses Phänomen tritt an späterer Stelle erneut auf, als L4 einen leistungsschwachen Schüler im Gegensatz zu seinen leistungsstärkeren Mitschüler*innen („zwei Drittel der Klasse“) mit einem fiktiven Namen bezeichnet („**oh** Johannes Egon hat Schwierigkeiten“). Während L3 seinen Kolleginnen zustimmt, führen L4 und L6 den Diskurs in einem generalisierenden Modus fort. Der Diskurs erreicht seinen Höhepunkt in L6s Aussage, mit der sie den bisherigen Diskussionsverlauf unter Verwendung von Parallelismen zusammenfasst (Z. 76–80). An dieser Stelle treten verdichtet emotionale Ausdrücke auf (Z. 75, 77). Es wird deutlich, dass leistungsschwache und -starke Schüler*innen mit Verhalten assoziiert werden, welches den Unterrichtsverlauf unterbricht und die Aufmerksamkeit der Lehrkräfte erfordert (Z. 76–78). Folglich bilden Schüler*innen der Mitte einen positiven Vergleichshorizont, da sie von den Lehrkräften als Beweis für die Korrektheit ihrer Leistungserwartungen herangezogen werden.

Das gegenseitige Verständnis der Lehrkräfte deutet auf einen gemeinsamen Orientierungsgehalt hin und dokumentiert sich auch auf formaler Ebene. L4s Eröffnung wird von zwei ihrer Kolleg*innen bekräftigt (Z. 73, 75), die daraufhin mit der Beschreibung der Mitte fortfahren. Alle vier Lehrkräfte sind dazu in der Lage abwechselnd auf „die Mitte“ Bezug zu nehmen, obwohl die Idee an keiner Stelle zuvor im Diskurs explizit eingeführt wurde. Auch weitere Spezifizierungen der Mitte führen nicht zu Konflikten. Da die Lehrkräfte ihre Äußerungen gegenseitig ergänzen und die Bedeutung der Äußerungen ihrer Kolleg*innen antizipieren, handelt es sich hierbei um einen parallelen Diskurs, der die Annahme strukturell äquivalenter Erfahrungen stärkt (vgl. Przyborski, 2004: S. 196 ff.).

4.2 Konzeptualisierungen von *underachievement* als (un-)sichtbar

Die in Form von Erzählungen skizzierten Schüler*innenbeispiele sind fundamental, um zu verstehen, wie UA durch die Lehrkräfte konstruiert und in Schüler*innen diagnostiziert wird. Der folgende Ausschnitt umfasst L2s Erfahrungen mit einem Schüler, bei dem sie UA vermutete:

64 **L2:** Jetzt gehts wiederum um=n anderes Beispiel wo=ich wiederum sehr wohl Einfluss
65 nehmen konnte, n Schüler der grundsätzlich großes Interesse an der
66 Fremdsprache::: gezeigt hat und auch (.) großes Potenzial aba insgesamt
67 verhaltensauffällig eh: war aba eher sag=ich mal so im::: eh::: aufgrund seines
68 @Alters@ und der Phase:: eh:::m um die es quasi ging (.) und da::: is=es mir schon
69 (2) hier und da mal erfolgreich gelungen und das auch im::: Fall dieses Schülers
70 ganz speziell (1) die Schüler hinter=m Ofen hervorzulocken

In ihrer Beschreibung des Schülers hebt L2 sein Interesse an der englischen Sprache, sein Potenzial und sein störendes Verhalten im Unterricht hervor. Ein Perspektivwechsel durch die formale Brille verdeutlicht, dass diese drei Merkmale in Verbindung zueinanderstehen: L2 schreibt ihrem Schüler ein Interesse am Fach zu und interpretiert diese

Annahme als notwendigen („grundsätzlich“) Indikator für sein beachtliches Potenzial. Sein störendes Verhalten hingegen scheint im Widerspruch („aba“) dazu zu stehen; es ist jedoch auffälliger und fungiert als Stigma („insgesamt verhaltensauffällig“).

Die Verhaltensauffälligkeiten des Schülers werden durch seine aktuelle Entwicklungsphase gerechtfertigt (Z. 67–68), die über einen bestimmten Zeitraum stabil sein wird. Trotz dessen nimmt sich L2 dieser Umstände an (Z. 64–65). Sie unterscheidet das Schülerbeispiel („jetzt gehts wiederum“) damit klar von dem Beispiel eines weiteren Schülers mit diagnostizierter Aufmerksamkeitsdefizit-Hyperaktivitätsstörung (ADHS), bei dem sie Hilflosigkeit verspürte (vgl. hierzu auch den Abschnitt über *twice-exceptionals* in Preckel & Baudson, 2013: S. 46–50).

Die Schilderungen von L5 sind strukturell ähnlich zu denen L2s, unterscheiden sich jedoch substantziell auf inhaltlicher Ebene. Im folgenden Ausschnitt reflektiert L5 die Herausforderungen des Unterrichts eines Schülers mit Erstsprache Englisch:

148 **L5:** Ich glaub ich hab vielleicht ja ok ehm::: ich hab zum Beispiel einmal auch n
 149 Muttersprachler im Unterricht gehabt das is halt auch immer noch so n bisschen
 150 schwierig weil man den ja auch irgendwie fordern will (.) ehm::: und trotzdem so=n
 151 bisschen also so zweigleisig zu fahren dass man immer gucken muss wo kann er
 152 sprachlich vielleicht was anderes machen (.) ehm::: weil er das andere schon
 153 draufhat und auf der anderen Seite is es ja trotzdem so dass wenn die jetzt
 154 zuhause als Muttersprache die ehm::: die eigentliche Fremdsprache schon
 155 können, dass die jetzt trotzdem aber ehm::: (.) auch gewissen Strategien oder so
 156 auch in der Fremdsprache fehlen und der hat jetzt in den Arbeiten ist der halt auch
 157 nicht super gut abgeschnitten weil er halt zum Teil auch ne ganz schlimme
 158 Rechtschreibung hatte

Die Ellipsen (Z. 148) zu Beginn von L5s Erzählung deuten auf eine gewisse Unsicherheit hin; es ist unklar, ob ihre Erzählung zu den Beispielen für UA passt, die ihre Kolleg*innen zuvor benannt haben. Darauf folgt erneut eine Beschreibung der Merkmale des Schülers (Z. 152–155). Strukturell identisch zu L2s Fallbeispiel verursacht die Tatsache, dass Englisch die Erstsprache des Schülers ist, hohe Erwartungen an sein Leistungspotenzial. Diese positive Bewertung steht im Kontrast (Z. 155) zu den Problemen des Schülers im Unterricht (Z. 157–158).

In diesem Fall verhindert die fehlerhafte Orthographie die Umsetzung des Potenzials des Schülers, das ihm durch L5 zugeschrieben wird. Da L5 die mögliche Diagnose von Lese-Rechtschreib-Schwierigkeiten (LRS) an späterer Stelle ausschließt und die fehlerhafte Rechtschreibung des Schülers stattdessen auf sein mangelndes Bemühen zurückführt, kann dieses als veränderbar betrachtet werden.² Dies führt zu einer ambigen Wahrnehmung des Schülers: Obwohl seine Erstsprache ihn grundlegend von seinen Mitschüler*innen unterscheidet, teilt er gleichzeitig die Probleme in der Orthographie mit ihnen. Diese Ambiguität spiegelt sich auch auf formaler Ebene wider, da das Unterrichten

² Dies impliziert nicht, dass jegliche Form der LRS unveränderbar ist. Ausgehend von einem inklusiven Verständnis von LRS (Gerlach, 2019: S. 12 f.) ist sogar das Gegenteil insbesondere dann der Fall, wenn die LRS durch äußere Umstände begründet sind. Mit Hilfe angemessener Förderung können Schüler*innen LRS in gewissem Maße kompensieren.

in der betroffenen Lerngruppe von L5 verlangt „zweigleisig zu fahren“. Später beschreibt sie diese Herausforderung als „schwierig da so=ne bisschen den Mittelweg zu finden“. L5 muss ständig aushandeln, in welcher Relation der betroffene Schüler zu seinen Mitschüler*innen steht, um ihn angemessen zu fördern.

UA wird in diesen zwei Fällen als *sichtbar* konzeptualisiert. Beide Lehrerinnen verweisen zunächst auf ein relativ statisches Merkmal ihrer Schüler (Interesse am Fach, Erstsprache Englisch), das sie positiv deuten. Dies führt dazu, dass die Leistungserwartungen der Lehrkräfte steigen. Die erhöhten Leistungserwartungen bleiben jedoch unerfüllt, was mit einem Verweis auf ein weiteres, negativ gedeutetes Merkmal der Schüler (störendes Verhalten, fehlerhafte Rechtschreibung) begründet wird. Das UA der Schüler ist für die Lehrkräfte sichtbar, weil es für sie aus dem Zusammenhang zweier beobachtbarer Schüler*innenmerkmale resultiert.

L3s „**Paradebeispiel**“ für UA bildet einen Kontrast zu den beiden zuvor geschilderten Fällen:

132 **L3:** Eh::: ja also ich hab eine Schülerin deren Eltern ehm::: (.) kein Deutsch sprechen
 133 (2) die also von zuhause aus (.) ehm::: nicht so=ne große Unterstützung erfahren
 134 kann (.) und ehm::: deren d-der Bildungshintergrund der Eltern is jetzt eh::: auch
 135 nich ehm::: so dass man das als ideale Voraussetzung ehm erstmal sehen würde
 136 und dieses Mädchen is aber richtig richtig clever und ehm::: (2) inner Schule nich
 137 immer ganz aufmerksam, und eh::: das is ne ganz eh::: fiese Kombination quasi
 138 zuhause nicht die nötige Unterstützung erhalten zu können und in der Schule nicht
 139 eh::: gut genug aufzupassen als das man das ohne Unterstützung alleine
 140 bewältigen könnte (.) und ehm::: ja diese Schülerin könnte deutlich mehr leisten
 141 als sie leistet ehm::: und die größte Aufgabe is eigentlich ihr das vor Augen zu
 142 führen (.) ehm möglichst einfühlsam dass sie zuhause diese Unterstützung nicht
 143 kriegen kann und dass sie aufpassen muss dass der Zug nicht eh ohne sie abfährt
 144 weil sie nämlich eigentlich sehr sehr clever is nur wenn der Zug einmal weg is
 145 dann:: is halt schwierig (.) das wär so mein meine **Paradebeispiel** eigentlich für
 146 *underachievement*

Erneut basiert die Erzählung auf den Merkmalen einer Schülerin. In diesem Fall erfolgt jedoch eine Erweiterung, da auch der familiäre Hintergrund der Schülerin berücksichtigt wird (Z. 132–136). Dabei ist L3s Redestil unsicher, vage und wenig strukturiert (Z. 132). Dies legt nahe, dass die Informationen über die Eltern der Schülerinnen einerseits relevant für die Legitimation ihrer Schulleistung sind, damit andererseits jedoch ein sensibles Thema angeführt wird. Trotz der herausfordernden Lebensumstände der Schülerin, schätze L3 sie als „richtig richtig clever“ ein.

L3 identifiziert die mangelhafte Aufmerksamkeit der Schülerin im Unterricht (Z. 136) als eine Ursache für ihr UA. Die häusliche Sphäre der Schülerin und die institutionelle Sphäre werden zu Komplementen (Z. 137). Hierdurch entsteht eine implizite Verbindung zwischen dem Verhalten der Schülerin im Unterricht und den Nachteilen, die für sie aufgrund des Bildungsgrades und der Deutschkenntnisse ihrer Eltern entstehen. L3 verwendet die Metapher eines losfahrenden Zuges (Z. 144), um zu erklären, wie er seine Schülerin unterstützt. Die Metapher deutet darauf hin, dass der Schülerin die Verantwortung für die Verbesserung ihrer Schulleistung zugeschrieben wird.

Es wird deutlich, dass die Irritation von L3s Leistungserwartungen an seine Schülerin gegensätzlich zu den beiden vorherigen Fällen konstruiert ist; UA ist hier *unsichtbar*. Die herausfordernden Lebensumstände der Schülerin bilden in Kombination mit ihrer geringen Aufmerksamkeit ein Merkmal, auf Grundlage dessen die Leistungserwartungen der Lehrkraft sinken. Die Schülerin gerät erst dadurch in L3s Fokus, dass er sie trotz ihres als defizitär konstruierten familiären Hintergrunds für „richtig clever“ hält. Während im Falle sichtbaren UA vereinzelte Merkmale der Schüler*innen die Umsetzung des von ihren Lehrkräften antizipierten Leistungspotenzials verhindern, erfolgt die Zuschreibung eines solchen Leistungspotenzials im Falle unsichtbaren UAs erst gar nicht.

Zusammenfassend lässt sich Folgendes beobachten: Alle drei Erzählungen beginnen mit einer – mehr oder weniger – detaillierten Beschreibung von Schüler*innenmerkmalen. Sie alle resultieren in besonders niedrigen (unsichtbares UA) oder hohen (sichtbares UA) Leistungserwartungen auf Seiten der Lehrkräfte. Durch die Betonung dieser Merkmale und Umstände wird die Einzigartigkeit der Schüler*innen im Vergleich zu ihren Mitschüler*innen unterstrichen. Die Tatsache, dass diese Aspekte von allen drei Lehrkräften direkt zu Beginn ihrer Erzählungen thematisiert werden, deutet darauf hin, dass sie besonders relevant für die Konzeptualisierung von UA sind. Demzufolge erfolgt die Identifikation von UA vielmehr aufgrund der Voraussetzungen der Schüler*innen anstatt aufgrund ihrer gezeigten Leistung. Schulleistung wird gemessen an dem Grad der Erfüllung der Erwartung der Lehrkräfte, die aus der Bewertung der Voraussetzungen ihrer Schüler*innen resultieren. Damit UA als Ursache für die schwachen Leistungen der Schüler*innen in Erwägung gezogen wird, müssen sie zunächst als *other* identifiziert werden.

4.3 Intervention bei *underachievement* im Modus der erzwungenen Kooperation

Nachdem die Lehrkräfte ihr Verständnis von UA ausgehandelt hatten, war es möglich, danach zu fragen, wie sie mit – ihres Erachtens nach – betroffenen Schüler*innen umgehen. Dabei wurde die erzwungene Kooperation als eine handlungsleitende Orientierung rekonstruiert. Diese umfasst alle Lehrkräfte, die grundsätzlich dazu bereit waren, ihre Schüler*innen dabei zu unterstützen, ihre Schulleistung zu verbessern. Diese Bereitschaft war jedoch an die Akzeptanz der Anweisungen der Lehrkraft durch die Schüler*innen gebunden und folglich erzwungen. Wie ersichtlich wird, hängt das Ergebnis dieses Modus der Intervention maßgeblich von dem Grad der Lehrer*innen-Schüler*innen-Beziehung ab. In der Gesamtschau zeigen die vier Fälle A–D, dass einige Lehrkräfte eine relativ enge Beziehung zu ihren Schüler*innen etablieren. Dies erlaubt ihnen, die individuellen Interessen ihrer Schüler*innen bei der Intervention zu berücksichtigen.

L1 arbeitete sechs Jahre lang mit einem Schüler, den sie als „wirklich clever“ jedoch „oft zu cool für den Unterricht“ beschreibt. Trotz ihres Bemühens, den Schüler zur Teilnahme am Unterricht zu motivieren, blieb er dem Unterricht fortgehend fern:

174 **L1:** Der war auch (2) ehm:: (3) den hatte ich auch sechs Jahre lang ganz cleveres
 175-178 Kerlchen (.) wirklich clever da hab=ich mir auch gedacht richtiges Potenzial (2) [...] **auf jeden Fall**
 179 hab=ich nich den Weg gefunden ihn (.) von seinem (.) cool sein
 180-184 @sag=ich jetzt mal@ [...] un::d der is tatsächlich bis zum (.) bis zum Schluss hinter
 185 seinen Leistungen geblieben (.) da war=ich auch noch ganz gemein @fällt mir
 186 grade ein@ ehm (.) hat ganz massiv dann auch geschwänzt (.) und ehm:: ich hab
 187 hinterher telefoniert was das Zeug hält ich hab auch gesagt wir entschuldigen nich
 188 mehr der Mutter die (.) das immer ne neue Ausrede hat sich einfallen lassen (.)
 189 und er hat tatsächlich auch nich von mir tatsächlich auch von mir nich den
 190 Abschluss bekomm den er verdient hätte also und da hat=er nachher kam er dann
 191 noch an nach n paar Monaten (2) hat mir erzählt er wäre bis jetzt jeden Tag in der
 192 Schule gewesen er wäre Klassensprecher und er hätte gute Noten (3) @ich sach
 193 dann ma weiter so@

In diesem Ausschnitt sind die L1s Maßnahme inhärente Autorität sowie der resultierende Widerstand des Schülers omnipräsent; die Spannung zwischen den Akteur*innen ist beinahe greifbar. Als Beweis der Qualität der Beziehung betont L1, der Schüler sei nach dem Schulwechsel von sich aus an sie herangetreten (Z. 190). Sie offenbart ihr Interesse an dem Lernerfolg des Schülers, indem sie metaphorisch die Rolle einer Mutterfigur einnimmt (vgl. Lange-Vester & Teiwes-Kügler, 2014: S. 190). Da die Mutter des Schülers in ihren Versuchen, ihren Sohn von einem regelmäßigen Schulbesuch zu überzeugen, scheitert, übernimmt L1 die Verantwortung (Z. 187). Dadurch erweitert sie ihren Handlungsspielraum über die Grenzen des Klassenzimmers hinaus. Trotz L1s Bemühungen ist es erst der Schulwechsel, der zu einer Verbesserung der Leistungen des Schülers führt (Z. 192).

Derselbe Modus tritt bei der Intervention in den weiteren Fällen auf. L4 machte beispielsweise Erfahrungen mit einer Schülerin, die aufgrund von Unterforderung wiederholt den Unterricht störte („die rückt da die Möbel grade“). Diese Störungen endeten erst durch die Vergabe zusätzlicher Aufgaben, also einer Art der Binnendifferenzierung („ich musste nich nur die irgendwie beschäftigen sondern der auch noch richtig was zu tun geben“). L2 berichtet von einer ähnlichen Situation, in der sie die Vorliebe ihres Schülers für Hip-Hop-Musik nutzt. Indem sie die Leidenschaft ihres Schülers in den Englischunterricht integriert („mich mit dem auszutauschen und so kleine Zusatzaufgaben“), verhilft sie ihm zu einer Leistungsverbesserung.

5. Implikationen für den Englischunterricht

Das Konzept der Mitte, welches in allen Fällen A–D diskutiert wurde, ist im Falle unerwartet niedriger Performanz relevant für die Unterrichtsgestaltung und Diagnostik der Lehrkräfte. Hinsichtlich der Unterrichtsgestaltung personifizieren die Schüler*innen der Mitte die impliziten Leistungserwartungen der Lehrkräfte. Solange die Mehrheit der Lerngruppe dem Unterricht problemlos folgt, erachten die Lehrkräfte das Niveau ihres Unterrichts für angemessen. Aus diesen Beobachtungen ergeben sich zwei Schlussfolgerungen: (1) Die Struktur des mehrgliedrigen Bildungssystems und die nicht mehr zeitgemäße Zielsetzung der Formung homogener Lerngruppen ist tief in der Praxis von Lehrkräften

verankert; (2) Die Forderung inklusiven Englischunterrichts ist noch lange nicht erfüllt, da alle Lehrkräfte (unbewusst) die Homogenisierung ihrer Schüler*innen verfolgen.

Gleichzeitig verdeutlicht die Signifikanz der Mitte, dass die Leistungsbewertung vor dem Hintergrund einer sozialen Bezugsnorm erfolgt. Dadurch ist eine niedrige Schulleistung stets in Relation zu der Lerngruppe und dem mit ihr einhergehenden durchschnittlichen Leistungsniveau definiert. Dies impliziert, dass Schüler*innen, die in einer leistungsstarken Lerngruppe unterdurchschnittliche Schulleistungen erzielen, in einer anderen Lerngruppe möglicherweise zur Mitte gehören. Um dieses Problem zu entschärfen, empfiehlt sich ein Rückgriff auf die individuelle oder kriteriale Bezugsnorm (vgl. Ingenkamp & Lissmann, 2008). Die individuelle Bezugsnorm erlaubt es Lehrkräften, die unterschiedlichen Lernverläufe ihrer Schüler*innen anzuerkennen und ist besonders im Hinblick auf inklusiven Englischunterricht erwähnenswert. Die kriteriale Bezugsnorm hingegen hilft bei der Kompensation eines Mangels an standardisierter Leistungserhebung, die aus dem Fokus der Lehrkräfte auf informelle Diagnostik hervorging.

Diese Beobachtung gleicht den Ergebnissen von Leucht et al. (2012), die empfehlen, bildungspolitische Dokumente wie den GER im Unterricht zu verwenden, um sicherzustellen, dass Lehrkräfte die norm- und kriterienorientierte Beurteilung voneinander trennen und dies auch bei der Unterrichtsgestaltung berücksichtigen. Die Eignung des GER als globaler Referenzrahmen für ein solches Vorhaben ist jedoch kritisch zu prüfen. Die resultierende Frage nach einem angemessenen Grad an Konkretion wird innerhalb des GER (2001) explizit thematisiert: „A common framework scale should be *context-free* in order to accommodate generalisable results from different specific contexts. [...] Yet at the same time the descriptors in a common Framework scale need to be *context-relevant*, relatable to or translatable into each and every relevant context“ (S. 21; Hervorhebung im Original). Somit könnte eine Aufgabe der Lehrkräftebildung zukünftig darin bestehen, die Deskriptoren des GER fachspezifisch zu konkretisieren und für die Leistungsbewertung nutzbar zu machen. In jedem Fall erscheint es legitim, dass die Abwendung von der sozialen Bezugsnorm das Risiko, die Schulleistung von Schüler*innen zu überschätzen (vgl. Südkamp et al., 2008; Zhu & Urhahne, 2015), verringert.

Weiterhin wurde deutlich, dass Lehrkräfte UA primär bei Schüler*innen vermuten, die sich von ihren Mitschüler*innen aufgrund auffällig niedriger Schulleistung und ihres Status als *other* unterscheiden. Da die Lehrkräfte UA nicht mit Blick auf Schüler*innen mit durchschnittlicher Schulleistung diskutierten, kann nicht ausgeschlossen werden, dass UA in dieser Gruppe unerkannt bleibt. Nahezu alle Schüler*innen, bei denen UA vermutet wurde, hatten gemeinsam, dass die Erwartung an ihr Potenzial aus einem statischen Merkmal resultierte. Eine solche Eigenschaft verursachte eine Generalisierung; die Identität dieser Schüler*innen war durch ihre Lehrkräfte als Agent*innen der sozialen Institution Schule determiniert, um es in Goffmans Worten zu sagen. Diese Generalisierung ist eng verbunden mit der Kategorisierung, die als eines der grundlegenden Prinzipien bei der Diagnostik fungierte („die war eben so n Typ Schüler“). Daher ist es möglich, dass Lehrkräfte relevante Eigenschaften ihrer Schüler*innen übersehen, sofern diese nicht den etablierten Kategorien entsprechen.

L3s Beispiel unsichtbaren UAs unterstreicht die Bedeutung von Habitussensibilität als essenzielle Voraussetzung und Kompetenz für die Integration marginalisierter Gruppen in Schule (vgl. Lange-Vester & Teiwes-Kügler, 2014). Indem L3 das Verhalten seiner Schülerin im Unterricht dem Mangel an Unterstützung im familiären Umfeld gegenüberstellt, beweist er „die Fähigkeit und Bereitschaft, sich gedanklich an den Ort zu versetzen, den ein Schüler oder eine Schülerin im sozialen Raum einnimmt“ (S. 177). Dies erlaubt ihm erst, den Einfluss der Erfahrungen der Schüler*innen auf ihr Verhalten im Englischunterricht zu reflektieren.

Die von den Lehrkräften beschriebenen Interventionsmaßnahmen fokussieren die Symptome des UAs anstelle der Ursachen. Obwohl die Interventionen in ihrer konkreten Ausgestaltung variieren, sind sie durch den Modus der erzwungenen Kooperation vereint. Die Schüler*innen sind nicht aktiv an der Aushandlung der Interventionsmaßnahme beteiligt; stattdessen erfolgt die Auswahl und Durchsetzung allein durch die jeweilige Lehrkraft. Somit besteht größtenteils keine Passung zwischen den Interventionen selbst und den Bedürfnissen der Schüler*innen, sodass eine Verbesserung der Schulleistung ausbleibt. Strukturell zusammengehalten werden die unterschiedlichen Interventionsmaßnahmen durch die ihnen inhärente Beziehungskomponente, die in der Forschung zu (Lehrer*innen-)Professionalität als pädagogisches Arbeitsbündnis konzeptualisiert wird. Nur wenige Lehrkräfte erhalten die Autonomie ihrer Schüler*innen während der Intervention aufrecht und erfüllen somit die Forderung Oevermanns (2002), „die Hilfe muss also zugleich eine zur Selbsthilfe sein“ (S. 26).

6. Schluss

Die rekonstruierten Orientierungen liefern erste, vage Informationen über die Identifikation von UA im Englischunterricht der Sekundarstufe. Sie deuten einerseits auf die Relevanz von pädagogischer Diagnostik und UA als Inhalte für die erste und zweite Phase der Lehrkräftebildung hin. Andererseits geben sie Aufschluss über die Bedeutung pädagogischer Haltung für im Englischunterricht eingesetzte diagnostische Verfahren und deren Güte. Um UA im Englischunterricht weiter zu explorieren, bedarf es zunächst einer fachspezifischen Schärfung des Konzepts der diagnostischen Kompetenz und weiterer Forschung zu der Entstehung diagnostischer Urteile (vgl. Leuders et al., 2020). Es ist weiter danach zu fragen, wie hoch der Anteil (unentdeckter) *underachievers* im Englischunterricht ist und welche pragmatischen und zugleich qualitätsvollen Verfahren Anwendung finden können, um UA zu identifizieren und betroffene Schüler*innen zu fördern.

Dennoch unterliegen die Ergebnisse einer Vielzahl von Einschränkungen. Erstens erfolgte die Rekonstruktion anhand des Datenmaterials einer höchst selektiven Auswahl an Englischlehrkräften der Sekundarstufe. Zudem nahmen keine Hauptschullehrkräfte an der Datenerhebung teil. Im Kontext von UA erscheint die Hauptschule, die sich in den vergangenen Jahren mit zunehmender Stigmatisierung konfrontiert sieht, als besonders relevante Schulform. Zweitens ist davon auszugehen, dass der Zeitmangel in einzelnen Erhebungssituationen (z. B. während einer zeitlich begrenzten Mittagspause) das

Redeverhalten der Lehrkräfte beeinflusste und sie zu relativ knappen Antworten verleitete. Die daraus resultierenden Nachfragen des Autors könnten Einfluss auf die Relevanzsetzungen der Lehrkräfte in ihren Erzählungen genommen haben. Drittens ist die Anzahl der Fälle unzureichend für eine Generalisierung der Ergebnisse im Rahmen der soziogenetischen oder relationalen Typenbildung (vgl. Nohl, 2013).

6. Literatur

- Baudson, Tanja G.; Fischbach, Antoine & Preckel, Franzis (2016). Teacher judgements as measures of children's cognitive ability: A multilevel analysis. In *Learning and Individual Differences*, 52, S. 148–156. <https://doi.org/10.1016/j.lindif.2014.06.001>
- Baumert, Jürgen & Kunter, Mareike (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. In *Zeitschrift für Erziehungswissenschaft*, 9, S. 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Berliner, David C. (1994). Expertise: The Wonders of Exemplary Performance. In Mangieri, John N. & Block, Cathy C. (Hg.), *Creating Powerful Thinking in Teachers and Students*. New York: Holt, Rinehart and Winston, S. 141–186.
- Bohnsack, Ralf (1989). *Generation, Milieu und Geschlecht – Ergebnisse aus Gruppendiskussionen mit Jugendlichen*. Opladen: Leske + Budrich.
- ___ (2000). Gruppendiskussion. In Flick, Uwe; Kardorff, Ernst von & Steinke, Ines (Hg.), *Qualitative Forschung: Ein Handbuch*. Berlin: Rowohlt, S. 369–384.
- ___ (2014). Documentary Method. In Flick, Uwe (Hg.), *The SAGE Handbook of Qualitative Data Analysis*. Thousand Oaks (CA): SAGE Publications, S. 217–233.
- Bohnsack, Ralf & Pfaff, Nicolle (2010). Die Dokumentarische Methode: Interpretation von Gruppendiskussionen und Interviews. In Maschke, Sabine & Stecher, Ludwig (Hg.), *Enzyklopädie Erziehungswissenschaft Online*. Weinheim: Beltz Juventa, n. p. <https://doi.org/10.3262/EE007100073>
- Bonnet, Andreas (2020). Die notwendige Zumutung der Komplexität und welche Früchte sie trägt. In *Zeitschrift für Rekonstruktive Fremdsprachenforschung* 1(1), S. 5–18. <http://www.rekonstruktive-fremdsprachenforschung.de/zeitschrift/ausgabe-1/Bonnet>
- Brühwiler, Christian; Helmke, Andreas & Schrader, Friedrich-Wilhelm (2017). Determinanten der Schulleistung. In Schweer, Martin (Hg.), *Lehrer-Schüler-Interaktion*. Wiesbaden: Springer VS, S. 291–314.
- Elsner, Daniela (2018). Institutionalised Foreign Language Learning—Teaching English at Different Levels. In Surkamp et al. (Hg.), S. 17–37.
- Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Flick, Uwe (2011). Das Episodische Interview. In Oelerich, Gertrud & Otto, Hans-Uwe (Hg.), *Empirische Forschung und Soziale Arbeit: Ein Studienbuch*. Berlin: Springer, S. 273–280.

- Gass, Susan M. & Selinker, Larry (2008). *Second Language Acquisition: An Introductory Course*. New York: Routledge.
- Gerlach, David (2019). *Lese-Rechtschreib-Schwierigkeiten (LRS) im Fremdsprachenunterricht*. Tübingen: Narr Francke Attempto.
- Gunia, Jürgen (2012). Kompetenz: Versuch einer genealogischen Ideologiekritik. In *Textpraxis*, 4(1), S. 1–11. <http://www.uni-muenster.de/textpraxis/juergen-gunia-kompetenz>
- Helmke, Andreas (2015). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts (6. Aufl.)*. Seelze-Velber: Friedrich Verlag.
- Helfferrich, Cornelia (2011). Interviewplanung und Intervieworganisation. In Helfferrich, Cornelia (Hg.), *Die Qualität qualitativer Daten*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 167–189.
- Ingenkamp, Karl-Heinz & Lissmann, Urban (2008). *Lehrbuch der pädagogischen Diagnostik (6. Aufl.)*. Weinheim: Beltz Juventa.
- Kersten, Kristin (2020). Der Zusammenhang von kognitiven Fähigkeiten und Zweisprachigkeit im frühen L2-Erwerb: Individuelle und externe Faktoren. In Böttger, Heiner; Festman, Julia & Müller, Tanja (Hg.), *Language Education and Acquisition Research*. Bad Heilbrunn: Klinkhardt, S. 82–116.
- KMK [Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland] (2012). *Bildungsstandards für die fortgeführte Fremdsprache (Englisch/Französisch) für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.10.2012*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_10_18-Bildungsstandards-Fortgef-FS-Abi.pdf (zuletzt aufgerufen am 13.10.2021)
- (2019). *Standards für die Lehrerbildung: Bildungswissenschaften*. **Fehler! Linkreferenz ungültig.** (zuletzt aufgerufen am 13.10.2021)
- Kolb, Annika (2011). Kontinuität und Brüche: Der Übergang von der Primar- zur Sekundarstufe im Englischunterricht aus der Perspektive der Lehrkräfte. In *Zeitschrift für Fremdsprachenforschung*, 22(2), S. 145–175.
- Kolb, Annika & Raith, Thomas (2018). Principles and Methods – Focus on Learners, Content and Tasks. In Surkamp et al. (Hg.), S. 73–88.
- Kubanek-German, Angelika & Edelenbos, Peter (2005). “Diagnostic competence” as a key qualification of the foreign language teacher in the 21st century. In Ahrens, Rüdiger (Hg.), *Englisch in der Erwachsenenbildung des 21. Jahrhunderts*. Heidelberg: Universitätsverlag Winter, S. 91–104.
- Küster, Lutz (2016). Kompetenzorientierung im Kontext des Lehrens und Lernens von Sprachen. In Burwitz-Melzer, Eva (Hg.), *Handbuch Fremdsprachenunterricht (6. Aufl.)*. Tübingen: A. Francke Verlag, S. 83–87.
- Lamprecht, Juliane (2007). *Die Bewertung von Schülerleistungen: Eine kritische Reflexion auf der Basis der dokumentarischen Evaluationsforschung*. Berlin: Logos Verlag.

- Lange-Vester, Andrea & Teiwes-Kügler, Christel (2014). Habitussensibilität im schulischen Alltag als Beitrag zur Integration ungleicher sozialer Gruppen. In Tobias Sander (Hg.), *Habitussensibilität*. Wiesbaden: Springer VS, S. 177–207.
- Leucht, Michael; Tiffin-Richards, Simon; Vock, Miriam; Pant, Hans Anand & Köller, Olaf (2012). Diagnostische Kompetenz von Englischlehrkräften. In *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(4), S. 163–177. <https://doi.org/10.1026/0049-8637/a000071>
- Leuders, Timo; Loibl, Katharina & Dörfler, Tobias (2020). Diagnostische Urteile von Lehrkräften erklären – Ein Rahmenmodell für kognitive Modellierungen und deren experimentelle Prüfung. In *Unterrichtswissenschaft* 48, S. 493–502. <https://doi.org/10.1007/s42010-020-00085-5>
- Lightbown, Patsy M. & Spada, Nina (2013). *How Languages are Learned (4. Aufl.)*. Oxford (GB): Oxford University Press.
- Loewen, Shawn (2020). *Introduction to Instructed Second Language Acquisition (2. Aufl.)*. New York: Routledge.
- Montgomery, Diane (2020). *Tackling Disadvantage and Underachievement in Schools: A Practical Guide for Teachers*. New York: Routledge.
- MSBNRW [Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen] (2011). https://www.schulentwicklung.nrw.de/lehrplaene/upload/lehrplaene_download/hauptschule/Engl_HS_KLP_Endfassung.pdf (zuletzt aufgerufen am 05.01.2022)
- (2019). *Kernlehrplan für die Sekundarstufe I Gymnasium in Nordrhein-Westfalen: Englisch*. https://www.schulentwicklung.nrw.de/lehrplaene/lehrplan/199/g9_e_klp_%203417_2019_06_23.pdf (zuletzt aufgerufen am 05.01.2022)
- (2021). *Masterplan Grundschule*. <https://www.schulministerium.nrw/system/files/media/document/file/Masterplan%20Grundschule.pdf> (zuletzt aufgerufen am 05.01.2022)
- Kemmerer, Alexandra (2019). Konzeption videobasierter und forschungsorientierter Lerneinheiten zur Förderung der Diagnosekompetenz von Lehramtsstudierenden im Fach Englisch. In Kreft, Annika & Hasenzahl, Mona (Hg.), *Aktuelle Tendenzen in der Fremdsprachendidaktik*. Berlin: Peter Lang, S. 63–78.
- Nohl, Arnd-Michael (2013). *Relationale Typenbildung und Mehrebenenvergleich*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Oevermann, Ulrich (2002). Professionalisierungsbedürftigkeit und Professionalisiertheit pädagogischen Handelns. In Kraul, Margret; Marotzki, Winfried & Schweppe, Cornelia (Hg.), *Biographie und Profession*. Bad Heilbrunn: Klinkhardt, S. 19–49.
- Porsch, Raphaela & Wilden, Eva. (2017). The Introduction of EFL in Primary Education: Challenges for EFL Teachers in Germany. In Wilden, Eva & Porsch, Raphaela (Hg.), *The Professional Development of Primary EFL Teachers: National and International Research*. Münster: Waxmann, S. 59–75.
- Preckel, Franzis & Baudson, Tanja G. (2013). *Hochbegabung: Erkennen, Verstehen, Fördern*. München: C. H. Beck.

- Przyborski, Aglaja (2004). *Gesprächsanalyse und dokumentarische Methode*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Przyborski, Aglaja & Wohlrab-Sahr, Monika (2014). *Qualitative Sozialforschung: Ein Arbeitsbuch (4. Aufl.)*. München: Oldenbourg Wissenschaftsverlag.
- Reichertz, Jo (1992). Beschreiben oder Zeigen – Über das Verfassen Ethnographischer Berichte. In *Soziale Welt* 43, S. 331–350.
- Rost, Detlef H. & Hanses, Petra (1997). Wer nichts leistet, ist nicht begabt?: Zur Identifikation hochbegabter Underachiever durch Lehrkräfte. In *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 29(2), S. 167–177.
- Schrader, Friedrich-Wilhelm & Helmke, Andreas (1987). Diagnostische Kompetenz von Lehrern: Komponenten und Wirkungen. In *Empirische Pädagogik* 1(1), S. 27–52.
- ___ (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In Weinert, Franz E. (Hg.), *Leistungsmessungen in Schulen*. Weinheim: Beltz, S. 45–58.
- Schrader, Friedrich-Wilhelm (2013). Diagnostische Kompetenz von Lehrpersonen. In *Beiträge zur Lehrerinnen- und Lehrerbildung* 31(2), S. 154–165.
- Siegle, Del (2018). Understanding Underachievement. In Pfeiffer, Steven I. (Hg.), *Handbook of Giftedness in Children*. New York: Springer Verlag US, S. 285–297.
- Sparfeldt, Jörn R., & Rost, Detlef H. (2012). Underachievement: Diskrepanz von Leistungspotential und schulischer Leistung. In *Erziehung & Unterricht* 162, S. 435–441.
- Spinath, Birgit (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer und das Konstrukt der diagnostischen Kompetenz. In *Zeitschrift für Pädagogische Psychologie* 19, S. 85–95.
- Stamm, Margrit (2011). Wenn Jugendliche ihr ‚Potenzial‘ nicht umsetzen – vier Perspektiven auf den Begriff Underachievement und seine Legitimationsprobleme. In Zlatkin-Troitschanskaia, Olga (Hg.), *Stationen Empirischer Bildungsforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 23–35.
- Südkamp, Anna; Möller, Jens & Pohlmann, Britta (2008). Der Simulierte Klassenraum: Eine experimentelle Untersuchung zur diagnostischen Kompetenz. In *Zeitschrift für Pädagogische Psychologie* 22(3–4), S. 261–276.
- Südkamp, Anna; Kaiser, Johanna & Möller, Jens (2012). Accuracy of teachers' judgements of students' academic achievement: A meta-analysis. In *Journal of Educational Psychology* 104(3), S. 743–762. <http://dx.doi.org/10.1037/0022-0663.99.2.253>
- Surkamp, Carola & Viebrock, Britta (Hg) (2018). *Teaching English as a Foreign Language*. Stuttgart: J.B. Metzler.
- Uhlig, Johannes; Solga, Heike & Schupp, Jürgen (2009). *Ungleiche Bildungschancen: Welche Rolle spielen underachievement und Persönlichkeitsstruktur?* WZB Discussion Paper, No. SP I 2009-503. Berlin: Wissenschaftszentrum Berlin für Sozialforschung.
- Valtin, Renate (2020). Eine Zwei ist eine Drei ist eine Vier: Über die mangelnde Vergleichbarkeit von Zensuren und die Problematik der Grundschul-Empfehlungen. In *Schule im Blickpunkt. Informationen des Landeselternbeirats Baden-Württemberg* 43(4), S. 6.

- Van Ackeren, Isabell; Klemm, Klaus & Kühn, Svenja M. (2015). *Entstehung, Struktur und Steuerung des deutschen Schulsystems: Eine Einführung (3. Aufl.)*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- VERBI Software. (2019). *MAXQDA 2020 [Computersoftware]*. Berlin: VERBI Software.
- Vygotsky, Lev S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge (MA): Harvard University Press.
- Weinert, Franz E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In Weinert, Franz E. (Hg.), *Leistungsmessungen in Schulen*. Weinheim: Beltz, S. 17–31.
- Wilden, Eva & Porsch, Raphaela (2020). A Critical Analysis of the German Public Debate about an Early Start in Primary Foreign Language Education. In *Language Teaching for Young Learners* 2(2), S. 192–212. <https://doi.org/10.1075/ltvl.19019.wil>
- Zhu, Mingjing & Urhahne, Detlef (2015). Teachers' judgements of students' foreign-language achievement. In *European Journal of Psychology of Education* 30, S. 21–39. <https://doi.org/10.1007/s10212-014-0225-6>

Über den Autor

Joel Guttke ist wissenschaftlicher Mitarbeiter am Lehrstuhl für Didaktik des Englischen an der Universität Duisburg-Essen. In seinem laufenden Dissertationsvorhaben entwickelt er ein fachspezifisches Erhebungsinstrument für das Potenzial zu kognitiver Aktivierung im Englischunterricht der Primarstufe. Des Weiteren ist er an inklusivem Englischunterricht (mit Schwerpunkt auf Schüler*innen entlang des Autismus-Spektrums) und dem Einsatz maschinellen Lernens im Englischunterricht interessiert.

Korrespondenzadresse: joel.guttke@uni-due.de