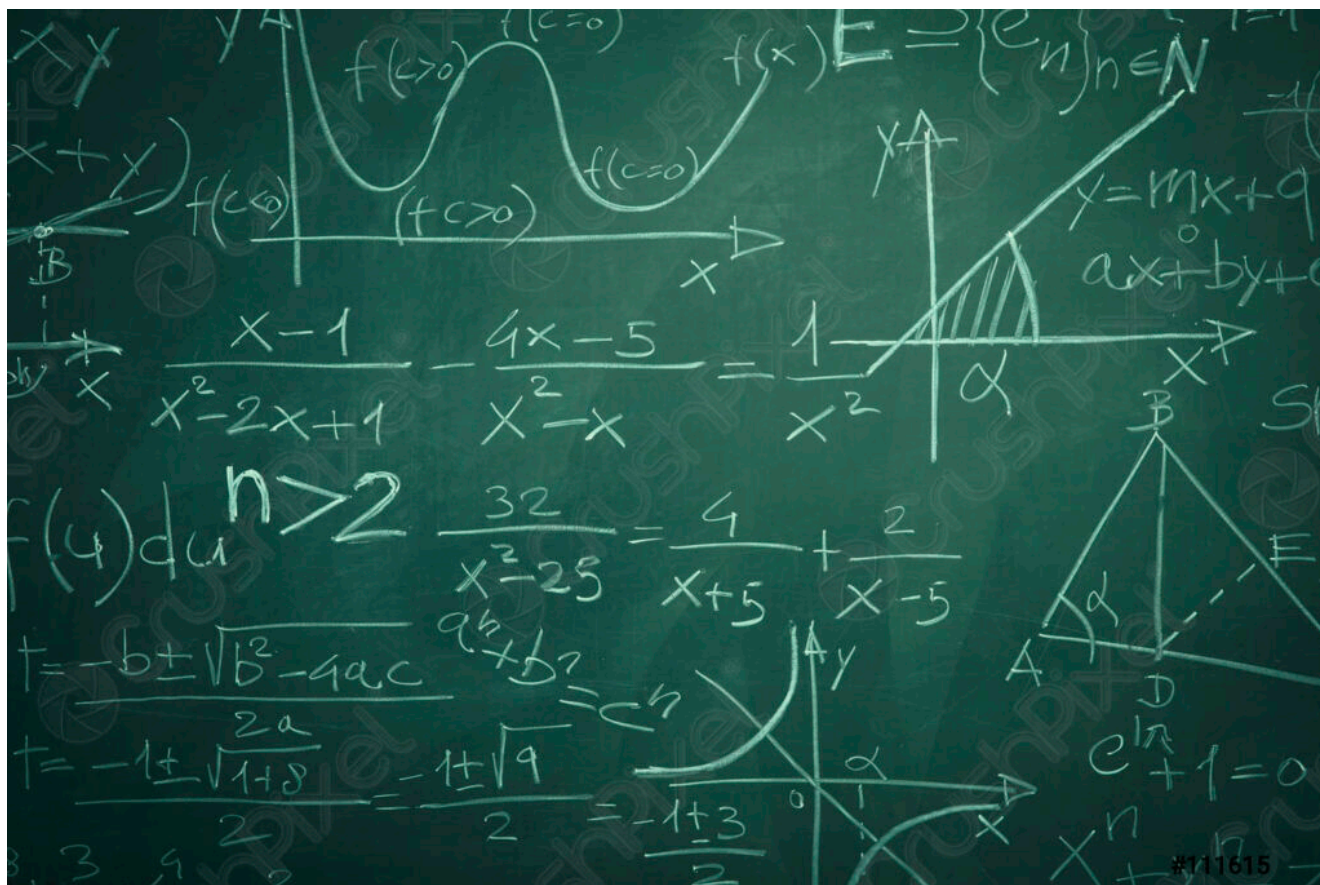


Zur inneren Logik mündlicher Prüfungen im Fach Mathematik: Teil 1

Category: Blog

geschrieben von Gastautor | 8. September 2022



Einführung

Im folgenden Beitrag soll die innere Logik mündlicher Prüfungen im Fach Mathematik mit dem Fokus auf die schweizerische Maturprüfung untersucht werden. Meine Ausführungen gründen dabei auf einer vor allem anekdotischen und weniger empirischen Basis. Doch auch wenn bekanntermassen «Daten» nicht der Plural von Anekdote ist, so ist die Stichprobengrösse, die meiner Analyse zugrunde liegt, nicht unerheblich: Seit elf Jahren unterrichte ich an einem Basler Gymnasium Mathematik und - in einem erheblich geringeren Umfang - Chemie. In dieser Zeit habe ich rund 250 Maturprüfungen (alle in Mathematik) abgenommen. Daneben bin ich seit sieben Jahren als Experte[1] für Mathematik an verschiedenen Gymnasien und Fachmaturitätsschulen (FMS) in

unterschiedlichen Schweizer Kantonen im Einsatz und habe damit etwa ebenso viele Prüfungen bei über einem Dutzend verschiedener Prüferinnen und Prüfer begleiten und protokollieren können. Hinzu kommt meine Tätigkeit an der Pädagogischen Hochschule Luzern, wo ich als Experte für Allgemeine Didaktik und Pädagogische Psychologie bei den Diplomprüfungen in den Bildungs- und Sozialwissenschaften in den letzten sechs Jahren rund 120 Prüfungen begleitet habe.

Es wird sich zeigen, dass ein solcher Versuch äusserst notwendig ist.

Zweifelsfrei gibt es Lehrpersonen und Experten bzw. Expertinnen, die in ihrer Berufskarriere mehr Prüfungen abgenommen bzw. begleitet haben. Gleichwohl scheinen mir 250 eigene und 250 beobachtete Maturitäts- und FMS-Prüfungen sowie 120 Diplomprüfungen eine ausreichende Basis für den Versuch zu sein, die innere Logik einer Abschlussprüfung im Fach Mathematik zu rekonstruieren und sie damit an den pädagogischen Ansprüchen, die an eine mündliche Prüfung gestellt werden, zu messen. Es wird sich zeigen, dass ein solcher Versuch äusserst notwendig ist.



Dr. Mario Gerwig, Jg. 1984, ist seit 2011 Lehrer für Mathematik und Chemie am Gymnasium

Leonhard Basel. Er ist Schulbuchautor und Länderberater (Mathematik Neue Wege, Westermann Schweiz) sowie Experte für Maturprüfungen in Mathematik (Kantone BS und BL) und Diplomprüfungen in den Bildungs- und Sozialwissenschaften (PH Luzern). 2021 erschien das Buch Der Satz des Pythagoras in 365 Beweisen (Springer Spektrum, Berlin).

Pädagogische und rechtliche Voraussetzungen einer mündlichen (Matur-)Prüfung

1. Gütekriterien mündlicher Prüfungen

Das anzustrebende Ideal einer jeden Leistungsüberprüfung sollte maximale Fairness sein, abzielen sollte sie auf die gerechte Beurteilung von Lernerfolgen. Das bedeutet, dass Prüfungen jeglicher Art die Leistung der Geprüften gültig, zuverlässig und objektiv abbilden sollten. Es geht also im Kern immer darum, Beurteilungsfehler zu vermeiden, was insbesondere bei mündlichen Prüfungen schwieriger umzusetzen ist, als es zunächst erscheinen mag. Denn Urteilsverzerrungen und Fehlschlüsse sind menschlich, der interaktive Echtzeitcharakter und der damit einhergehende Handlungsdruck sind eine besondere Herausforderung. Gleichzeitig liegt die Vermeidung mancher Fehler nicht in der Hand des Prüfers, was insb. individuelle Faktoren der Beurteilten wie den Einfluss der Tagesform oder mögliche externe Störfaktoren im Prüfungsumfeld betrifft. Andere Fehler hingegen sind vermeidbar.

Als Prüfer und Prüferin muss man sich der möglichen Fehlerquellen bewusst sein, um sie vermeiden zu können.

Von besonderer Bedeutung bei mündlichen Prüfungen sind (vgl. Bohl 2009) der *Sozialgruppeneffekt* (Sprache der Aufgabenstellung wird von bestimmten Personengruppen schlechter verstanden), der *Nähe-Fehler* (Prüfungen nach einer schwachen Prüfung werden tendenziell zu gut bewertet und andersherum), der *Reihungsfehler* (mehrere Prüfungen nacheinander dienen als Referenzpunkte, obwohl sie unabhängig voneinander beurteilt werden sollten) und der *Wissen-um-die-Folgen-Fehler* (wenn die möglichen Konsequenzen einer Bewertung bekannt sind, bspw. das Nichtbestehen der Maturität bzw. des Abiturs, wird eine Prüfung oft milder beurteilt). Als Prüfer und Prüferin muss man sich der möglichen Fehlerquellen bewusst sein, um sie vermeiden zu können. Dazu ist es hilfreich und sinnvoll, sich an den Gütekriterien der *Validität*, *Reliabilität* und *Objektivität* einer Leistungsbeurteilung zu orientieren, denn die Wahrscheinlichkeit sinkt, einen der genannten Fehler zu begehen, je besser diese eingehalten werden. Die Erfahrungsberichte im folgenden Kapitel zeigen, zu welchen Auswüchsen es in mündlichen Prüfungen kommen kann, wenn gegen eines oder mehrere der Gütekriterien in beträchtlichem Masse verstossen wird.



Prüfungen jeglicher Art sollten die Leistung der Geprüften gültig, zuverlässig und objektiv abbilden.

Die *Validität* einer Prüfung ist dann gegeben, wenn das, was überprüft werden soll, auch tatsächlich überprüft wird. Dies ist deshalb eine besondere Herausforderung, weil etwa eine bestimmte Fähigkeit meist nicht direkt zu beobachten ist. Soll bspw. überprüft werden, ob ein Schüler das Urnenmodell der Kombinatorik verstanden hat, müssen Annahmen darüber getroffen werden, woran sich dieses „Verstehen“ festmachen lässt. Das Anwenden einer der vier Fälle des Modells auf ein bestimmtes, vorgegebenes Beispiel (vgl. den ersten Erfahrungsbericht des folgenden Abschnitts), vermag eine solche Feststellung sicherlich noch nicht rechtfertigen. Gelingt es dem Schüler hingegen, das

Konzept mit eigenen Worten zu erläutern und an einem Beispiel zu illustrieren, kann eher davon ausgegangen werden, dass das Konzept weitreichend verstanden ist. Entscheidend für die Validität sind daher gut begründete Annahmen, in diesem Fall läge der Bewertung die Annahme zugrunde, dass ein Sachverhalt nur dann mit eigenen Worten wiedergegeben werden kann, wenn er auch verstanden wurde. Wird hingegen eine einzige Beispielrechnung korrekt notiert, kann davon noch nicht zwangsläufig ausgegangen werden. Wir werden bei den im folgenden Abschnitt dargestellten Erfahrungsberichten sehen, dass insb. die Validität einer Prüfung häufig ein Problem darstellt.

Die Prüfung muss stets sorgfältig zusammengesetzt sein. Aufgaben mit einem mittleren Schwierigkeitsindex ergeben in der Regel eine höhere Reliabilität der Prüfung, daher empfiehlt es sich, nicht eine zu grosse Anzahl Aufgaben hoher Schwierigkeit einzubauen.



In einer mündlichen Prüfung sollte es vor allem darum gehen, im Prüfungsgespräch auf die Äusserungen des Prüflings einzugehen und einen bestimmten Sachverhalt zu diskutieren.

Die *Reliabilität* einer Prüfung ist dann gegeben, wenn ein Messinstrument, bspw. eine Aufgabe, bei wiederholter Anwendung dieselben Resultate liefert. Dieses Gütekriterium ist nicht ohne weiteres einzuhalten, denn dazu müssten eigentlich die eingesetzten Aufgaben im Vorfeld der Prüfung untersucht werden, was statistisch aufwendig und für eine Lehrperson im Rahmen einer mündlichen Abschlussprüfung schlichtweg nicht zu leisten ist. Aus theoretischen

Überlegungen und empirischen Untersuchungen ergeben sich jedoch einige Heuristiken, mit deren Hilfe eine mündliche Prüfung möglichst reliabel erstellt und durchgeführt werden kann (vgl. Lienert/Raatz 1998): Die Prüfung muss stets sorgfältig zusammengesetzt sein. Aufgaben mit einem mittleren Schwierigkeitsindex ergeben in der Regel eine höhere Reliabilität der Prüfung, daher empfiehlt es sich, nicht eine zu grosse Anzahl Aufgaben hoher Schwierigkeit einzubauen. Zufallsabhängige Aufgaben, insb. Multiple-Choice-Aufgaben mit nur zwei Antwortalternativen, vermindern die Reliabilität. In einer mündlichen Prüfung sollte es nun aber vor allem darum gehen, im Prüfungsgespräch auf die Äusserungen des Prüflings einzugehen und einen bestimmten Sachverhalt zu diskutieren. Um die Reliabilität zu steigern ist es dennoch sinnvoll, zentrale Fragen, welche das Prüfungsgespräch leiten und strukturieren, schriftlich vorzubereiten. Am wirksamsten verbessert man die Reliabilität über die Länge einer Prüfung, denn durch das Hinzufügen von Fragen kann diese erhöht werden. Dabei kann im Sinne der Ökonomie einer Prüfung diese natürlich nicht unbegrenzt verlängert werden, zumal diesbzgl. strikte Rahmenbedingungen einzuhalten sind. Doch wenn bspw. eine Prüfung schon bei einer ersten, eher einfachen Einstiegsaufgabe stecken bleibt, sind grosse Zweifel bzgl. der Reliabilität (und damit auch bzgl. der Validität) angezeigt, auch wenn diese Anfangsaufgabe am Ende korrekt gelöst werden sollte (vgl. den zweiten Erfahrungsbericht im folgenden Kapitel).

Eine nicht-objektive Prüfung ist niemals reliabel oder valide, eine nicht-reliable Prüfung kann objektiv sein, ist aber niemals valide. Gleichzeitig ist eine objektive und reliable Prüfung nicht zwangsläufig valide.

Die *Objektivität* einer Prüfung ist dann gegeben, wenn verschiedene Massnahmen angewendet werden, um den Prüfungsprozess zu standardisieren und damit Ursachen für Messfehler zu reduzieren. Dazu gehört bspw. das Herstellen einer Atmosphäre, die eine konzentrierte und gleichzeitig entspannte Prüfung ermöglicht, das Einrichten einer Sitzordnung, bei der man nicht zu weit auseinander- aber auch nicht zu nah beieinandersitzt sowie eine ruhige und optimistische Ansprache, wenn man die Prüflinge begrüsst, sie in den Raum hineinbittet und das Prozedere der Prüfung erläutert. Die Objektivität leidet deutlich, wenn (vgl. den ersten Erfahrungsbericht im folgenden Kapitel) der Prüfling kaum begrüsst und stattdessen schon die erste Frage gestellt bekommt, wenn dieser noch nicht einmal seinen Platz eingenommen hat. Natürlich steht in

einem solchen Fall auch die Validität infrage, da ein solcher Beginn zu Prüfungsstress führt und der Prüfling seine eigentliche Leistung möglicherweise nicht mehr abrufen kann. Zur Steigerung der Objektivität ist es zudem empfehlenswert, die Leistungsmessung von der Leistungsbeurteilung zu trennen. Nach einer Prüfung könnten dazu anhand eines Kriterienrasters zunächst Punkte vergeben werden, bevor darauf basierend eine Note ermittelt wird. Die Objektivität steigt zudem deutlich, wenn die Prüfung möglichst von einer externen Fachperson protokolliert wird, die ihrerseits ebenfalls die Prüfung beurteilt (in der Schweiz fällt diese Aufgabe dem Experten bzw. der Expertin zu). So können etwaige Unterschiede im Anschluss diskutiert und eine gemeinsame Bewertung unter Einbezug des Protokolls argumentativ ausgehandelt werden.



Die Gütekriterien der klassischen Testtheorie können ein sinnvolles Orientierungsraster sein, um eine Prüfung möglichst fair und gerecht zu gestalten.

Zusammenfassend lässt sich feststellen, dass die Gütekriterien der klassischen Testtheorie ein sinnvolles Orientierungsraster sein können, um eine Prüfung möglichst fair und gerecht zu gestalten, d.h. um eine gültige, zuverlässige und objektive Leistungsmessung zu realisieren. Die Validität einer Prüfung ist dabei primäres Ziel. Für diese ist die Reliabilität der Beurteilung eine notwendige, aber nicht hinreichende Bedingung. Objektivität schliesslich ist eine notwendige, aber ebenfalls nicht hinreichende Bedingung für eine reliable Prüfung. Mit anderen Worten: Eine nicht-objektive Prüfung ist niemals reliabel oder valide, eine nicht-reliable Prüfung kann objektiv sein, ist aber niemals valide. Gleichzeitig ist eine objektive und reliable Prüfung nicht zwangsläufig valide. Es ist daher entscheidend, das Zusammenspiel der Kriterien zu beachten und für die

möglichst gute Einhaltung alle drei Gütekriterien zu sorgen.

2. EDK und KMK

Die rechtlichen Grundlagen für die schulischen Abschlussprüfungen in Deutschland und der Schweiz verweisen ihrerseits nicht auf die oben genannten Gütekriterien. Sie legen kaum inhaltliche Massstäbe fest, definieren dafür unterschiedlich streng die einzuhaltenden Rahmenbedingungen.

Die etwaige Einrichtung einer Vorbereitungszeit direkt vor der mündlichen Prüfung ist nicht einheitlich geregelt und kann zwischen den Kantonen, sogar zwischen einzelnen Schulen und Fächern und bisweilen sogar innerhalb einer Fachgruppe variieren.

In der Schweiz wird dies in der durch den schweizerischen Bundesrat in Kraft gesetzte „Verordnung über die schweizerische Maturitätsprüfung“ vom 7. Dezember 1998 geregelt. Sie legt u.a. die Zuständigkeiten fest (Art. 2), skizziert den Prüfungszweck (Art. 8), definiert Richtlinien (Art. 10) und bestimmt die Prüfungsfächer (Art. 14). Über die Prüfungsart informiert Art. 18, in dem es heisst: „In den Grundlagenfächern (...) und Mathematik (...) wird schriftlich und mündlich geprüft“ (EDK 1998, S. 1420). Mündliche Mathematikprüfungen im Rahmen der schweizerischen Maturität sind somit obligatorisch. Die konkrete Ausgestaltung wird jeweils von den Kantonen festgelegt, in der Regel dauern die schriftlichen Prüfungen zwischen drei und vier Stunden, die mündlichen Prüfungen mindestens 15 und höchstens 20 Minuten. Die etwaige Einrichtung einer Vorbereitungszeit direkt vor der mündlichen Prüfung ist nicht einheitlich geregelt und kann zwischen den Kantonen, sogar zwischen einzelnen Schulen und Fächern und bisweilen sogar innerhalb einer Fachgruppe variieren. Ist eine Vorbereitungszeit in aller Regel bei Prüfungsfächern wie Deutsch, Spanisch oder Französisch eine Selbstverständlichkeit, so ist es andersherum ebenso selbstverständlich, dass die mündlichen Maturprüfungen in Mathematik ohne eine solche abzulegen sind. Es sei angemerkt, dass es schwierig sein dürfte, eine *pädagogische* Begründung für diesen Umstand vorzubringen.

In Deutschland regelt die „Vereinbarung über die Abiturprüfung der gymnasialen Oberstufe in der Sekundarstufe II“ die Organisation der Abiturprüfungen. Auch

hier wird u.a. ein Rahmen für die mündlichen (§7) und schriftlichen Prüfungen (§5) sowie deren Korrektur, Beurteilung und Bewertung definiert (§6), der von den einzelnen Bundesländern auszugestalten und von den Schulen umzusetzen ist. Interessant ist die Tatsache, dass für alle mündlichen Prüfungen, die im Rahmen der Abiturprüfung abgelegt werden müssen, eine Vorbereitungszeit fest vorgeschrieben wird. In §7.5 heisst es: „Die Aufgabenstellung einschliesslich der Texte wird dem Prüfling schriftlich vorgelegt. Während der Vorbereitung unter Aufsicht darf sich der Prüfling Aufzeichnungen machen. Die Vorbereitungszeit beträgt in der Regel 20 Minuten.“ (KMK 2008, S. 5)

[1] Ein externer Experte bzw. eine externe Expertin ist in der Schweiz für alle mündlichen Maturprüfungen vorgeschrieben. Er bzw. sie protokolliert die Prüfung, achtet auf ein angemessenes Niveau sowie eine Vergleichbarkeit der Bewertungen innerhalb einer Klasse. Die Expertentätigkeit wird im Art. 12.2 der „Verordnung über die schweizerische Maturitätsprüfung“ vom 7. Dezember 1998 (geändert am 22. April 2009) definiert: „Die Experten und Expertinnen nehmen an den mündlichen Prüfungen in den verschiedenen Fächern teil und nehmen Einsicht in die schriftlichen Arbeiten. Anhand der Leistungen in den schriftlichen und mündlichen Prüfungen nehmen sie eine Gesamtbeurteilung der Kandidaten und Kandidatinnen vor“ (EDK 1998, S. 1750). Mit der zweiten Teilprüfung sind die Maturprüfungen in den fünf Prüfungsfächern gemeint, zur ersten Teilprüfung gehören die übrigen Fächer, welche ohne Maturprüfung abgeschlossen werden, deren Noten aber ebenfalls im Maturzeugnis aufgeführt werden.