



PH



«Teaching to the Test» bei Lernpass plus?

Eine Untersuchung zum Zusammenhang zwischen
der Durchführung von Orientierungstests und
Leistungen in den Stellwerktests

Prof. Dr. Michael Kickmeier-Rust
Institut Kompetenzdiagnostik
Pädagogische Hochschule St.Gallen

März 2023

EINFÜHRUNG

Das Thema "teaching to the test" (TtT) wird im Kontext von Lernpass plus und Stellwerktest immer wieder von verschiedenen Seiten und aus verschiedenen Perspektiven diskutiert. In vielen Fällen stellt sich bei genauerer Betrachtung heraus, dass der Verdacht auf TtT lediglich auf Missverständnissen über die Funktionsweise von Lernpass plus beruht. Es gibt aber auch legitime Bedenken, insbesondere zur Struktur des Aufgabenpools, die eine genauere Untersuchung des Themas rechtfertigen. Durch die PHSG wurde daher im Herbst 2022 eine Analyse durchgeführt, um zu untersuchen, ob TtT im aktuellen Lernfördersystem möglich ist und sich durch Daten belegen lässt. In der Analyse der Daten der Fächer Mathematik und Deutsch für die 8. Klasse aus dem Schuljahr 2021/22 wurde untersucht, ob sich Hinweise für eine gezielte oder exzessive Vorbereitung der SuS auf den Stellwerktest in der Praxis finden lassen. Dabei wurden die Bearbeitungszahlen sowie die Zusammenhänge zwischen der Durchführung von Orientierungstests und den Stellwerk-Resultaten betrachtet.

GRUNDLAGEN: FUNKTIONSBEREICHE UND AUFGABENPOOLS BEI LERNPASS PLUS

Lernpass plus umfasst zwei Bereiche: den standardisierten, normativen Stellwerktest (SW) und den Bereich Fördern, der sich wiederum aus Orientierungstests (OT) und Übungsaufgaben zusammensetzt. Die Orientierungstests dienen der Ermittlung des aktuellen Lernstands, die eigentliche Förderung geschieht durch das Arbeiten mit den Übungsaufgaben.

Lernpass plus arbeitet mit zwei getrennten Aufgabenpools: Ein Pool für die Testaufgaben und ein Pool für die Übungsaufgaben. Bei den Übungsaufgaben erhalten die Lernenden jeweils eine Rückmeldung zu der korrekten Lösung. Bei den Testaufgaben (SW und OT) erhalten die Lernenden keine Hinweise, ob die gewählte Lösung korrekt oder falsch war. Sowohl Stellwerk als auch die Orientierungstests sind adaptive Tests. Adaptiv bedeutet, dass ein Algorithmus aus dem Pool von mehreren hundert Testaufgaben für alle SuS individuell die Aufgaben auswählt, die jeweils den meisten Informationsgehalt innerhalb des individuellen Testverlaufs haben. Alle SuS erhalten daher unterschiedliche Aufgaben bzw. Aufgabensequenzen.

DEFINITION "TEACHING TO THE TEST"

Zunächst muss der Begriff TtT eingegrenzt werden: TtT bedeutet eine in irgendeiner Weise «unstatthafte», gezielte bzw. eine zu verengte Vorbereitung auf einen bestimmten definierbaren Test, die das Erzielen eines möglichst guten Resultats im Test ermöglicht, ohne dass damit ein allgemeiner Leistungszuwachs im getesteten Bereich verbunden ist. Das bedeutet, dass prinzipiell die konkreten Inhalte bzw. auch die konkreten Aufgaben gezielt im Unterricht thematisiert und geübt werden. Ein solches Vorgehen beinhaltet u.a. die Information über konkrete Inhalte, die Vermittlung konkreter Inhalte, oder die Präsentation «echter» Testaufgaben (vgl. Popham, 2001).

Im Kontext von Lernpass plus und Stellwerk würde TtT also bedeuten, dass die SuS durch das Arbeiten mit OT oder Förderaufgaben die konkreten Testaufgaben des SW und deren Lösungen testspezifisch lernen könnten, sodass sie in die Lage versetzt würden, gute Stellwerk-Ergebnisse zu erzielen, ohne gleichzeitig einen generellen Leistungszuwachs aufzuweisen.

Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16-2.

"TEACHING TO THE TEST" DURCH FÖRDERN IN LERNPASS PLUS?

Die Grundannahme des Verdachts auf TttT ist, dass es durch die Verwendung des Förderbereichs in Lernpass plus in irgendeiner Weise zu TttT in Bezug auf den Stellwerttest kommen könnte.

Da die Förderaufgaben wie oben beschrieben mit einem gänzlich separierten Aufgabenpool arbeiten, ist ein TttT für den SW hier von vornherein ausgeschlossen. Aufgrund des gemeinsamen Aufgabenpools von OT und SW wäre es aber theoretisch denkbar, dass es durch die Durchführung von OT zu TttT kommen könnte.

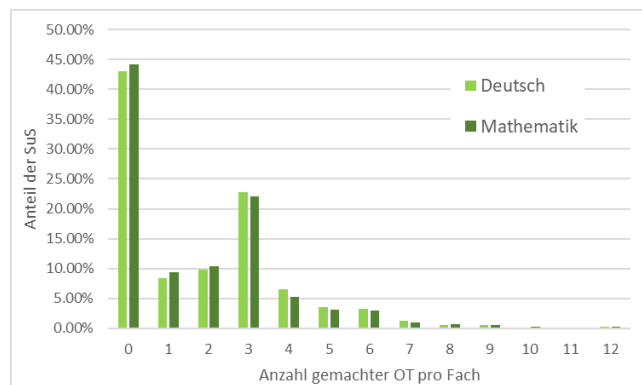
WERDEN OT EXZESSIV ALS VORBEREITUNG AUF DEN SW VERWENDET?

Ein häufig geäussertes Verdacht ist, dass sich die SuS durch exzessive Nutzung der OT auf den SW vorbereiten könnten. Durch wiederholtes Einstudieren der Aufgaben der OT würden sie sich so Kenntnis über den Aufgabenpool des SW verschaffen.

Diese Annahme scheint auf den ersten Blick aufgrund des gemeinsamen Testpools plausibel, allerdings lässt schon allein die Grösse des Testpools eine solche Strategie als nicht sehr erfolgsversprechend erscheinen: Im System kann jeder OT nur maximal 4-mal je Fach absolviert werden und umfasst jeweils knapp über 20 Aufgaben. Der Testpool beinhaltet pro Fach hunderte Aufgaben. Es bräuchte also eine sehr hohe Anzahl an OT-Durchführungen, um den gesamten Pool eines Fachs kennenzulernen.

Betrachtet man die Daten, zeigt sich, dass in der Praxis die Häufigkeit der OT-Durchführung sogar noch weit tiefer liegt: Die deutliche Mehrheit der SuS, die einen SW-Test gemacht haben, haben zuvor keinen OT gemacht. Wenn OT gemacht wurden, dann in der Regel 3 OT, also typischerweise 1 OT pro Kompetenzbereich eines Faches.

Anhand der Daten lässt sich also feststellen, dass es keinerlei Hinweise darauf gibt, dass in der Praxis eine exzessive Nutzung der OT als Vorbereitung auf den SW-Test stattfindet.



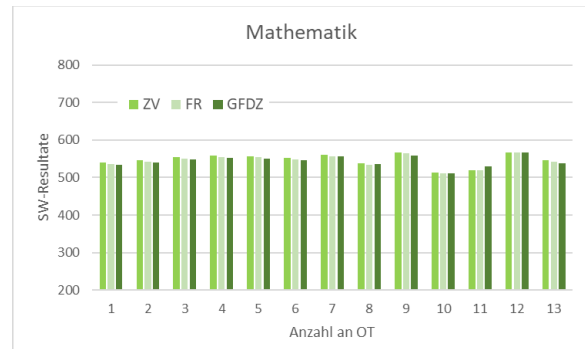
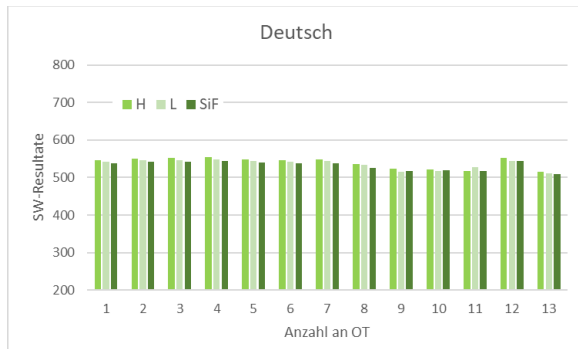
Anzahl an OT, die SuS vor einem SW-Test absolviert haben, getrennt nach den Fächern Deutsch und Mathematik.

BRINGT DIE VERMEHRTE DURCHFÜHRUNG VON OT EINEN VORTEIL FÜR DEN SW?

Wie oben festgestellt, gibt es keine weit verbreitete Praxis, exzessiv OT zu lösen. Es stellt sich aber die Frage, ob diejenigen SuS, welche eine grössere Anzahl an OT lösen, gegenüber denjenigen SuS, welche weniger OT lösen, einen Vorteil im SW haben. Wenn dies so wäre, würden die SW-Ergebnisse der SuS besser, je mehr OT sie zuvor absolviert hätten.

Wie die Daten belegen, hat die Anzahl gemachter OT keinen Einfluss auf die SW-Resultate. Die folgenden Abbildungen zeigen die SW-Resultate (y-Achse) in Bezug auf die insgesamt gemachten OT (pro Kompetenzbereich können maximal 4, also insgesamt 12 OT gemacht werden).

Es lässt sich sagen, dass die OT zwar ein guter Prädiktor für die Leistung in SW sind (es besteht eine Korrelation von 0.85 in Mathematik und 0.74 in Deutsch), die bloße Durchführung von einem oder mehreren OT führt aber nicht zu einer Verbesserung der SW-Resultate. Gute SuS zeigen sowohl in OT als auch in SW gute Leistungen. Letztlich spricht dieser Zusammenhang für die Qualität der Testdiagnostik in den LFS.



SW-Resultate, getrennt nach Kompetenzbereichen, in Bezug auf die Anzahl der zuvor in dem Fach insgesamt (also in allen Kompetenzbereichen) absolvierten OT.
[H ... Hören, L ... Lesen, SiF ... Sprache im Fokus; ZV ... Zahl und Variable, FR ... Form und Raum, GFDZ ... Grössen, Funktionen, Daten, Zufall]

WIE GROSS IST DIE ÜBERLAPPUNG DER AUFGABEN VON OT UND SW?

Anhand der Daten wurde genauer analysiert, wie oft es in der Praxis tatsächlich zu einer Überlappung bei einzelnen Aufgaben in OT und SW kommt, also wie viele der über 30 Aufgaben eines SW die SuS bereits zuvor in einem OT bearbeitet hatten. Die Untersuchung ergab, dass SuS, die mindestens einen OT gemacht hatten, im Schnitt etwa 5 Aufgaben eines Kompetenzbereichs schon in einem OT gesehen haben. Die folgende Tabelle stellt die Überlappung zwischen OT und SW dar.

Deutsch				
	H	L	SiF	Gesamt
Median	4.5	1.5	2.5	8.5
Mittelwert	4.07	3.16	3.69	10.94
SD	2.78	1.98	2.12	5.82
Mathematik				
	ZV	FR	GFDZ	Gesamt
Median	6	4	2	13
Mittelwert	6.23	4.68	3.45	14.36
SD	3.89	3.62	3.18	8.46

In den Daten zeigt sich also, dass es teilweise zu Überlappungen bei den Aufgaben von OT und SW kommt, auch wenn in der Regel der grösste Teil des Stellwerttests aus Aufgaben besteht, den die SuS zuvor noch nie in einem OT gesehen haben.

VORAUSSETZUNGEN FÜR "TEACHING TO THE TEST"?

Wie sich gezeigt hat, erhalten die SuS im SW also zu einem kleineren Teil Aufgaben, die sie bereits zuvor in einem OT bearbeitet haben. Es stellt sich die Frage, ob aufgrund dieses Sachverhalts bereits ein TttT zu bejahen wäre. Damit es beim SW durch die Verwendung von OT zu TttT kommt, müsste aber neben der vorherigen Vorlage einer Aufgabe auch gegeben sein, dass die Aufgaben und deren Lösungen den SuS zum Zeitpunkt des SW bekannt sind. Nur so könnten sie sich testspezifisch auf den SW vorbereiten.

Die SuS müssten dazu als erstes alle Aufgaben ihrer OT bis zur Durchführung des SW präzise in Erinnerung behalten. Da zwischen Durchführung der OT und der Durchführung des SW oft mehrere Monate liegen, ist schon dies nicht leichthin anzunehmen. Zweitens müssten die SuS auch die korrekten Antworten aller Aufgaben kennen und diese für den SW memorieren. Im OT selbst erhalten die SuS aber kein Feedback, was die richtige Lösung einer Aufgabe ist, sodass die Bearbeitung einer Frage allein noch keinen Vorteil bringt. Damit die Lösungen aller potentiellen SW-Aufgaben gelernt werden könnten, müssten die SuS diese nach jeder OT-Durchführung aus der Erinnerung heraus mit beträchtlichem Anstrengungen selber recherchieren. Der Aufwand hierfür wäre wohl kaum kleiner, als sich die Lerninhalte direkt aus dem Unterricht anzueignen. Und selbst wenn ein Lernender diesen Aufwand auf sich nähme, wäre damit noch kein gutes SW-Ergebnis garantiert, da nur ein kleinerer Teil des Testpools in Erfahrung gebracht worden wäre.

POTENTIELL DENKBARES, ECHTES "TEACHING TO THE TEST"

Echtes TttT, also eine tatsächlich effektive und gezielte Vorbereitung der SuS auf SW mittels Kenntnis der Testaufgaben und deren Lösungen, wäre nur dann möglich, wenn man es schaffen würde, sich über fast den gesamten Testpool Kenntnis zu verschaffen und die Testaufgaben auswendig zu lernen.

Dies ist eigentlich nur bei missbräuchlichem Verhalten denkbar. Dafür müssten unter enormem Aufwand in jedem Fach Hunderte Aufgaben mit zweifelhafter Intension kopiert und deren Lösungen recherchiert werden.

Durch die SuS wäre dies kaum möglich: Selbst bei maximaler Ausnutzung der möglichen OT (4 pro Teilbereich) ist es nicht möglich, den ganzen Testpool kennenzulernen. Ausserdem sind SuS im Normalfall auch nicht in der Lage die Aufgaben der OT zu kopieren, insofern die Lehrpersonen ihre Aufsichtspflicht während der Tests korrekt erfüllen.

Echtes TttT wäre daher eigentlich nur dann denkbar, wenn das unstatthafte Kopieren der Aufgaben durch eine missbräuchlich agierende Lehrperson geschehen würde. Falls solche Fälle von Fehlverhalten in der Praxis tatsächlich vorkommen sollten, läge das eigentliche Problem hier aber nicht beim System, sondern am missbräuchlichen Verhalten der betreffenden Lehrperson.

FAZIT UND WEITERES VORGEHEN

Aufgrund der Untersuchung lässt sich festhalten, dass es in der Praxis keine Indizien für ein TttT gibt: Weder werden OT exzessiv für die Vorbereitung auf den SW eingesetzt, noch finden sich in den Daten Belege dafür, dass die vermehrte Durchführung von OT den SuS einen messbaren Vorteil im SW verschafft. Aufgrund der gemeinsamen Nutzung des Testpools in SW und den OT kommt es allerdings gegenwärtig zu einem gewissen Grad an Überlappung der Testfragen, sodass SuS im SW teilweise Aufgaben erhalten, welche sie bereits in einem OT bearbeitet haben. Da die Aufgaben des Orientierungstests aber keine Rückschlüsse auf die richtigen Antworten erlauben, verschafft eine solche Überlappung im Grundsatz den SuS noch keinen Vorteil, selbst wenn sie eine Aufgabe aus einem früheren OT wiedererkennen sollten.

Während sich in der Studie zwar keine Hinweise auf tatsächlich existierendes TttT finden, erscheint die teilweise Überlappung der Fragen von SW und OT aber dennoch als nicht ideal. Um in diesem Punkt das System weiter zu verbessern und das Vertrauen in die Aussagekraft des Stellwerktests zu erhöhen, wurden mehrere Massnahmen beschlossen: Neben der bereits laufenden konsequenten Erweiterung des Aufgabenpools, sollen langfristig die Aufgabenpools für Orientierungstests und Stellwerk komplett getrennt werden. Als Sofortmassnahme wird der aktuelle Testalgorithmus mit dem Ziel angepasst, dass der Stellwerktest, wo immer möglich, Fragen ausfiltert, welche bereits in einem Orientierungstest vorgelegt wurden. Auf diese Weise soll eine Überlappung der verwendeten Aufgaben von Orientierungstests und Stellwerk verhindert werden.