F.1

- liebe Prüfungsvorsitzende, Frau Schäfer-Koch, liebe Prüfer:innen, Samuel, Cora und Augustin, liebe Kolleg:innen und liebe Studierende

- ich freue mich, Ihnen heute mit diesem Disputationsvortrag Einblicke in meine Dissertation bieten zu dürfen zum Thema Teachers' engagement with educational science how to communicate ...

- zusammengefasst geht es nun also um die nutzerfreundliche Kommunikation bildungswissenschaftlicher Ergebnisse an Lehrpersonen

F.2

- zu Beginn möchte ich die Relevanz des Themas herausstellen, bevor ich näher auf die entsprechenden wissenschaftlichen Diskurse der EISP und WK eingehe

- anschließend gebe ich ihnen Einblicke in die 3 experimentellen Studien meiner Dissertation

- ich schließe ganz klassisch mit einer Diskussion

F.3

- Wenn wir im Haus uns über bildungswissenschaftliche Erkenntnisse informieren, lesen wir - so vermute ich - vor allem wissenschaftliche Originalpublikationen zu Primär- und Sekundärforschung

F.4

- Aber auch in den Medien z.B. in Zeitungsartikel oder in Pressemitteilungen wird über bildungswissenschaftliche Ergebnisse informiert

F. 5

- Oder in Formaten, die gezielt Lehrpersonen adressieren

- Podcasts und Clearinghouses fassen bildungswissenschaftliche Forschungsergebnisse explizit für Lehrpersonen zusammen

F. 6

- Aber warum explizt für Lehrpersonen? Wie Sie sicherlich alle wissen werden Lehrpersonen formal <!-- z.B. in den KMK Kompetenzen in DE --> dazu angehalten solche Ergebnisse in ihrem professionellen Handeln zu berücksichtigen

- der Einbezug empirischer Studien und Theorien der BiWi wird seit Ende der 1990 unter verschiedenen Schlagwörtern diskutiert - beispielsweise unter dem Begriff ...

F. 7

- evidenzinformierte Praxis

- noch näher erläutern warum evidenzinformierte Praxis gefordert wird

F.8

- die Forderung evidenzinformiert Schulpraxis kann beispielsweise kognitionspsychologisch begründet werden

- Evidenz ist im Vergleich zu bespielsweise eigenen Erfahrungen als objektivere Informationsquelle anzusehen

- Da die Information systematisch nach bestimmten Kriterien gewonnen wird und dadurch weniger von kognitive Verzerrungen beeinflusst ist

- Ihr wird daher Potenzial zugeschrieben

- Professionalisierung von Lehrpersonen stärken

- Unterrichts- und Schulqualität

- letztlich Schüler:innenleistung steigern kann

- was auch mit ersten empirischen Studien belegt werden kann

- - Um die Potenziale zu entfalten, kann Evidenz Evidenz ganz unterschiedlich in der Praxis genutzt werden

- Evidenz z.B. zu Forschendem Lernen kann die eigene Unterrichtspraxis innovieren

- Evidenz kann aber auch genutzt werden, um Tipps für Schüler:innen abzuleiten z.B. wie Sie am besten Vokabeln lernen

- welche Schritte müssen Lehrpersonen durchlaufen, um solche praktischen Schlussfolgerungen aus der Evidenz abzuleiten?

F. 9

- angelehnt an das Prozessmodell von Groß Ophoff umfasst evidenzinformierte Praxis idealtypisch fünf Phasen

- konzeptuelle Phase, in der ein Anlass formuliert wird

- Zugriffsphase - hier wird zum Anlass passende wissenschaftliche Evidenz recherchiert

- Schlussfolgerungsphase: hier werden die wissenschaftlichen Ergebnisse rezipiert und in eigenes Vorwissen, Erfahrungen und Überzeugungen integriert, um Wissen zu generieren

- Anwendungsphase: wird das Wissen in die Praxis transferiert

- Phase 5 idealerweise wird der Prozess dann noch evaluiert,

F.10

- wie sie sicherlich auch wisse: gibt es eine Kluft zwischen diesen theoretischen Implikationen und der tatsächlichen Praxis

- Lehrpersonen greifen kaum auf Evidenz in ihrem professionellen Handeln zurück

- was nicht nur auf ihre Kompetenzen und Überzeugungen zurückzuführen ist, sondern empirisch zeigen sich vielzählige Barrieren, die unterschiedlichen Ebenen zugeordnet werden können

- neben der Ebene der individuellen Lehrperson

- Ebene der Forschungsergebnis. Ein Beispiel: Forschungsergebnisse sind durch die Verwendung zahlreicher Fachbegriffe z.B. statistischer Fachtermini schwer verständlich

- Schule

- Kommunikation: Kooperation zwischen Wissenschaftler:innen und Praktiker:innen im Forschungsprozess

F.11

- in der Literatur finden sich mittlerweile unterschiedliche Ansätze um Barrieren in Gelingensbedingungen umzuwandeln

- Diese können wieder in die vier Ebenen, die ich gerade genannt habe, unterteilt werden oder noch übergreifender in zwei Ansätze gegliedert werden

- Professionalisierung der Lehrpersonen ansetzen, indem gezielt Kompetenzen, die für die unterschiedlichen Phasen notwendig sind, gefördert werden

- Gestaltung der WK

- Zugänglichkeit von Evidenz steigern,

- Evidenz nutzerfreundlich z.B. verständlicher aufzubereiten

F.12

- in meiner Diss auf letzteren Ansatz aus unterschiedlichen Gründen fokussiert

- immer mehr WK-Formate vorhanden

- aber wenig Forschung wie Lehrpersonen mit diesen umgehen und gleichzeitig bietet dieser Bereich besonderes Potenzial, um Studien zu designen, die intern valide Erkenntnisse erlauben

- ressourcensparender als z.B. die Etablierung langfristiger RP Partnerships

- WK ist ein...

F.13

* sehr breites sowie interdisziplinäres FeldZusammenfassend: Aufbereitung nicht evidenzbasiert, wir wissen also nicht, ob wirklich leichter verständlich, informativer und praktisch relevant für Lehrpersonen
* Daher…

F.14

- wie zu Beginn angedeutet, ganz verschiedene WK-Formate vorhanden

- nun auf CHs beispielhaft eingehen, da im wissenschaftlich Diskurs zu EIP immer prominenter werden

F.15

- verschiedene Beispiele; vermutlich bekanntestes CH in DE: CHU

- Alle CHs sogenante knowledge brokers, die darauf abzielen Wissen zwischen zwei Communities – Praktiker:innen hier Lehrpersonen und Wissenschaft– zu vermitteln

- setzten unterschiedliche Schwerpunkte, aber fassen alle \*\*quantitative Studien\*\* zusammen, die sie als hochwertig eingestuft haben

- schaut man sich die Zusammenfassungen der quantitativen Studien näher an, zeigt sich, dass die statistische Ergebnisse aber ganz unterschiedlich aufbereitet werden

F.16

- Bspe. mitgebracht

- es ist zu erkennen dass textliche, aber auch visuelle Aufbereitungsformate genutzt werden

- und sowohl über Inferenzstatistiken - wie die klassische Signifikanzaussage -

- und Effektstärken - wie die in der Wissenschaft gängige standardisierte Effektstärke Hedges g informiert wird

- \*\*Visualisierung öffnen\*\* es finden sich aber auch der Theorie nach intuitivere Effektstärken wie Cohen's U3, die in Prozent den Anteil der einen Gruppe angibt, der über dem MW der anderen Gruppe liegt

- Zusammenfassungen bieten also Informationen darüber, wie sicher ein Effekt ist und wie Groß dieser ausfällt und somit die Grundlage um praktische Relevanz aus den Ergebnissen abzuleiten

- eingangs erläutert, dass CH darauf abzielen, wissenschaftliche Informationen verständlicher aufzubereiten, stellt sich aber die Frage, ob Lehrpersonen diese ganz verschiedenen Aufbereitungen tatsächlich korrekt interpretieren

- Aber warum ist diese Frage im Kontext evidenzinformierter Praxis so relevant?

F.17

- wenn wir uns nochmals an das Prozessmodell erinnern, dann führen falsche Interpretationen dazu,...

F.18

- dass die Schlussfolgerungsphase misslingt

- Das bedeutet, dass in der Schlussfolgerungsphase falsches Wissen generiert wir, dass dann entweder...

F.19

- zu einer inadäquaten Anwendungsphase führt oder...

F.20

- zu keiner Anwendungsphase

- denn die wissenschaftliche Evidenz wird abgewertet, sie wird nicht als praktisch relevant angesehen und der Prozess wird gestoppt

F.21

\*\*rhetorisches Highlight\*\* - daher: korrekte Interpretation statistischer Infos eine \*\*notwendige\*\*, aber \*\*nicht hinreichende\*\* Bedingung, um evidenzinformiertes Wissen in die Praxis adäquat oder überhaupt zu transferieren und letztendlich das Potenzial des ganzen Prozesses zu entfalten

- daher in meiner Dissertation den Fokus hierauf gelegt

- trotz dieser Relevanz gibt es kaum Forschung in BiWi wie Lehrpersonen solche Inferenzstatistiken und Effektstärken verstehen und wahrnehmen

- zwar erste Studien z.B. aus der Psychologie,

- die vor allem in experimentellen Studien untersuchen,

- wie wissenschaftliche Laien im allgemeinen diese statistische Informationen verstehen und wahrnehmen

- aber Forschungsstand sehr eklektisch z.B. U3 nie über mehrere Studien hinweg mit den gleichen Alternativen verglichen, was die Vergleichbarkeit der Studienergebnisse erschwert

- unklar welche Aufbereitungen von Inferenzstatistiken und Effektstärken besonders verständlich sind und sich für solche WK-Formate wie CHs eignen

- Daher…

F.22

- habe ich mich folgendem Forschungsinteresse gewidmet:

- wie Inferenzstatistiken und Effektstärken an Lehrpersonen textbasiert und visuell kommuniziert werden können, damit diese

- die Evidenz als praktisch relevant erachten,

- sie gleichzeitig die statistische Informationen akkurat interpretieren

- aber auch ihre kognitiven Ressourcen geschont werden

- nimmt man Akkuratheit und kognitive Ressourcen zusammen, bildet dies das Ziel der leicht verständlichen WK

- Dieser Aspekt ist der Fokus im Folgenden

F.23

- drei experimentelle Studien vorstellen, die jeweils Teile des Forschungsinteresses adressieren,

- aber jetzt je Studie nur eine ausgewählte FF, die sich zudem nur auf den Aspekt 'leicht verständliche statistische Informationen' bezieht

F.24

- kommen wir zur ersten Studie und zur FF wie Lehrpersonen statistische Beschreibungen in PM interpretieren

F.25

- hierfür repräsentativ 414 Lehrpersonen in DE rekrutiert

- in einem unvollständig rotieren 5x6 within within Design Lehrpersonen 2 PM präsentiert, die hinsichtlich des Within-Faktors Thema und des Within Faktors statistische Beschreibung manipuliert waren

- UV z.B. U3

F.26

- als AV zum einen Interpretation als Inferenzstatistik

- Likert-Item nach der Zustimmung gefragt, inwiefern die statistische Beschreibung eine Aussage darüber erlaubt, wie sicher es ist, dass der Unterschied tatsächlich exisitert - Interpretation als Inferenzstatistik -

- und zum anderen Interpretation als Effektstärke

- ähnliche Operationalisierung aber mit der Frage, ob statistische Beschreibung eine Aussage darüber erlaubt, wie groß der Unterschied ist

F.27

- kommen wir zu den Ergebnissen

- mit Blick auf die FF wie Lehrpersonen statistische Beschreibungen interpretieren herausfinden, dass Lehrpersonen nicht zwischen Inferenzstatistik und Effektstärke differenzieren

- näher an der Visualisierung erläutern

F. 28

- ein zentrale Punkt, den ich mit dieser Visualisierung veranschaulichen möchte ist, dass bei allen statistischen Beschreibungen die Verteilungen der Interpretationen als Inf und Eff stark überlappen, Interpretation als Inf links und der Interpretation als Eff rechts

- also...

F.29

- die Mittelwerte beider Interpretationen liegen nah beieinander \*\*in Anführungszeichen sagen\*\*

- obwohl eigentlich - außer bei der Kombi - die Beschreibungen entweder eine Inferenzstatistik oder ein Effektstärkemaß abbilden

F. 30

- korrekte Interpretation würde daher also so aussehen: Mittelwerte weit auseinander,\*\*Anführungszeichen\*\* genauer Verteilungen dürften kaum überlappen

- mithilfe einer Bayesianischen messwiederholten ANOVA konnten wir starke Evidenz für diese fehlende Differenzierung gewinnen was übrigens nicht unserer präregistrierten Hyp. entsprach

F. 31

- ich möchte noch auf ein anderes interessantes Ergebnis aufmerksam machen

- schauen wir uns die signifikanz und Cohen's U3 Beschreibung nochmals näher an

F.32

- die Interpretation der Signifikanz-Aussage als Effektstärke ist bimodal verteilt

- gleiches gilt für Interpretation von Cohen's U3 als Inferenzstatistik

- das lässt vermuten, dass es unterschiedliche Interpretationsgruppen gibt.

- Um ein Beispiel zu geben: eine Gruppe an Lehrpersonen gibt, die sich sicher ist, dass die Signifikanz-Aussage keine Aussage über die Größe eines Unterschieds erlaubt, aber es gibt auch eine Gruppe, die sich fälschlicherweise sicher ist, dass diese Aussage zulässig ist.

- Der Frage, ob solche unterschiedlichen Interpretationsgruppen vorliegen, wollte ich näher auf den Grund gehen, also haben Kolleg:innen und ich ...

F. 33

- in einer weiteren Studie diese Daten re-analysiert, um herauszufinden, ob Lehrpersonen verschiedenen Interpretationsprofile zugeordnet werden können

F.34

- da diese Studie eine Reanalyse war, haben wir auf die Daten aus Studie 1 zu Signifikanz und Cohen's U3 zurückgegriffen

F.35

- Durch das Explorieren mit latenten Profilanalysen ergaben sich verschiedene Interpretationsprofile, die überwiegend über beide Stichproben hinweg vergleichbar waren

- manche Lehrpersonen waren durchaus in der Lage die Signifikanz oder U3 korrekt zu interpretieren, hier in grün abgetragen und als Disentanglers bezeichnet

- aber Sie sehen in den rot dargestellten Profilen, dass die Mehrheit der Lehrpersonen Profilen zugeordnet werden können, die unterschiedliche Fehlkonzepte zusammenfassen

kommen wir jetzt noch zur 3 Studie

F.36

- nun visuelle Aufbereitung von Effektstärken gewidmet und der Frage nachgegangen, inwieweit Anreicherungen in Visualisierungen Akkuratheit der Interpretationen steigern

F.37

220 englischsprachige Lehrpersonen in einem 4x10x2x4 Between-Within-Within-Between Design jeweils insgesamt 20 Effektstärkenvisualisierungen präsentiert

- \_\_der Between-Faktor Thema kontextualisierte diese Effektstärkenvisualisierungen \_\_

- die wiederum in der präsentierte Effektstärke

- und in der Art der Visualisierung

- sogenannte Halfeye plots

- wir haben diese präsentiert, da wir in vorherigen Studien, die auch Teil dieses Projektes waren, diese als besonders geeignet für die Kommunikation mit Lehrpersonen identifizieren konnten

- zentrale UV Experimentalbedingungen Between-Faktor, also ob Anreicherung vorhanden oder nicht

- zwei von vier näher vorstellen

- Kontrollgruppe: halfeye plots ohne Anreicherung

- Signaling U3: Unterschiede im Sinne von Cohen's U3 hervorgehoben, also der Anteil der einen Gruppe über dem MW der anderen Gruppe lag

F.38

- Akkuratheit operationalisiert als

- abstrakte Metrik und U3 Metrik

- Lehrpersonen sollten die Größe der visuell präsentierte Effektstärke zum einen

- auf einer Skala von -1 komplette Unterlegenheit einer Gruppe bis +1 komplette Überlegenheit der Gruppe angeben

- und in % im Sinne der U3 Metrik angeben

- dummy Variable U3-Fehlkonzepte gebildet

- Personen, die im Median unplausibel kleine U3 Werte angaben wurden als Personen mit U3-Fehlkonzept definiert

F.39

- Ergebnisse basieren auf baysianischen Mehrebenenmodellen

- hinsichtlich Akkuratheit Evidenz für die Nullhypothese

- es gab also keine Unterschiede zwischen Visualisierungen mit Signaling U3 und ohne

- interessant ist aber, dass es Unterschiede zwischen den Operationalisierungen gab

- es gelingt Lehrpersonen erstaunlich gut, die visuell präsentierte Effektstärke in abstrakte Metrik zu transferieren

- sieht man in dieser Visualisierung gut \_\_erst jetzt Viz öffnen\_\_, \_\_erst jetzt öffnen\_\_ auf x-Achse die Differenz aus geschätzter und wahrer Effektstärke abgetragen

- in allen Bedingungen, Mittelwerterte um 0 tummeln, 0 = bedeutet hier die perfekte Übereinstimmung; außerdem moderate Streuung

- gleiche Grafik, aber jetzt U3 Ratings dargestellt: Verzerrungen in der U3 Metrik deutlich. Es liegt ein Bias vor, weil MW von 0 abweichen (Ergebnisse ohne Personen mit U3 Fehlkonzpet); Streuung moderat, die U3-Einschätzungen sind also im Durschnitt nicht akkurat

- positiv hervorzuheben ist, dass wir starke Evidenz dafür finden, dass Signaling U3 U3-Fehlkonzepte deutlich reduziert

F.40

zentralen Ergebnisse aller drei Studien nochmals in Take Home Messages zusammenbringen

F.41

\*\*langsamer!\*\*

- der Theorie nach intuitivere statistische Informationen...

- Signaling-Methode

F.42

- abschließend nochmals hervorheben, dass die Interpretation von statistischen Informationen zwar notwendig, aber nicht hinreichend für gelingendes evidenzinformiertes Handeln

- sie ist nur ein Faktor, der z.B. neben Überzeugungen, die Schlussfolgerung beeinflusst

- und insgesamt ist die Schlussfolgerungsphase wiederum nur ein kleines Puzzleteil eines komplexen Prozesses

F.43

- Meine Ergebnisse zeigen, dass die Gestaltung verständlicher WK herausfordernd

- weitere Forschung notwendig, um herauszufinden welche textlichen und visuellen Formate sich hierfür besonders eignen.

- jetzt erlaube ich mir aber abschließend noch etwas herauszuzoomen

- vier Jahre mit diesem Thema nutzerfreundliche WK beschäftigt und mir auch immer wieder Gedanken über das größere Ganze gemacht

- dabei bin ich auf ein Modell von Kriz und Hegarty zu Verstehensprozessen bei Animationen gestoßen, das mich auf die folgende Implikation für meine Studien gebracht hat

- diesem Modell nach, haben wir mit unserem Ansatz sogenannte bottom-up Verstehensprozesse adressiert

- wir haben versucht mithilfe von bestimmten Präsentationsformaten Lehrpersonen die Schritte hin zu einem adäquaten mentalen Modell zu erleichtern

- und letztendlich Lernprozesse auszulösen, die bei Lehrpersonen zu langfristigem Wissen führen, das für evidenzinformiertes Handeln notwendig ist

- das Modell impliziert aber auch, ...

F.44

- ... dass die Schritte und das Mentale Modell selbst durch das bereits vorhandene Wissen einer Lehrperson beeinflusst wird - was als top-down Verstehensprozesse bezeichnet wird

- angepasst an meine Forschungs könnten diese top-down Verstehensprozesse durch die statistische Kompetenz und die graph literacy einer Lehrperson beeinflusst sein

- Erste Hinweise für dieses Zusammenspiel auch in den Daten gezeigt

- um ein Beispiel zu geben

- in Studie 3 wiesen manche Lehrpersonen in der einen Akkuratheitsoperationalisierung ein Fehlkonzept auf, aber nicht in einer anderen, in keiner oder in mehreren Operationalisierungen, was möglicherweise auf die Ausprägung ihrer statistischen Vorkenntnisse zu Effektstärken zurückzuführen ist

- zusammenfassend also festhalten, dass WK alleine möglicherweise nicht ausreichend ist

- In zukünftigen Ansätzen sollte stärker das Zusammenspiel der Ansätze Wissenschaftskommunikation und Lehrpersonen Aus- und Fortbildung, also die systematische Förderung der Kompetenzen von Lehrpersonen berücksichtigt werden, um beide Verstehensprozesse zu adressieren, das Verstehen statistischer Informationen zu erleichtern, und einer gelingenden evidenzinformierten Praxis zumindest einen Schritt näher zu kommen