

Empirische Operationalisierung der Prinzipien von Meta-Reflexivität

Andreas Schmidt, Sebastian Röhl und Colin Cramer

Abstract

Dieser Beitrag macht die professionstheoretisch postulierten Prinzipien einer meta-reflexiven Lehrer:innenbildung für empirische Professionsforschung anschlussfähig. Dazu werden die meta-reflexiven Prinzipien schrittweise operationalisiert und in vier Pilotierungsstudien unter Lehramtsstudierenden validiert. Im Ergebnis liegt ein Instrument mit 18 Items vor, welches empirisch drei Dimensionen der Bedeutungszuschreibung an meta-reflexive Prinzipien in der Lehrer:innenbildung abbildet, die mit Vorläufigkeit, Multiperspektivität und Erkenntnisgenese bezeichnet werden.

Mit der Perspektive der Meta-Reflexivität wurde eine bislang ausschließlich theoriebasiert postulierte Perspektive auf Professionalität von Lehrpersonen in den professionstheoretischen Diskurs zum Lehrer:innenberuf eingebracht (Cramer et al., 2019). Anders als etwa der strukturtheoretische, kompetenzorientierte oder (berufs-)biografische Ansatz von Professionalität, die (auch) empirische Befunde zum Ausgangspunkt der Theoriebildung heranziehen oder ihre Theorie aus der Empirie ableiten (vgl. Cramer, 2020), steht eine empirische Durchdringung von Meta-Reflexivität noch aus.

Zwar haben Cramer, Groß Ophoff und Schreiber (2023) Überzeugungen von Lehrpersonen zu unterschiedlichen professionstheoretischen Ansätzen operationalisiert, darunter auch Meta-Reflexivität, aber dieser bislang singuläre empirische Zugang bezieht sich auf die subjektive Zuschreibung von Merkmalen an eine professionelle Lehrperson. Auch lag der Fokus dort auf einer Erfassung von Überzeugungen zu Professionalität insgesamt, nicht auf einer konsequenten Operationalisierung von Meta-Reflexivität bezogen auf seine vielfältigen theoretisch postulierten Dimensionen. Schon zuvor haben Hartmann et al. (2021) die Theoriefigur der Meta-Reflexivität für empirische Forschung genutzt, allerdings nicht um Meta-Reflexivität selbst zu untersuchen, sondern indem sie das theoretisch postulierte Prozessmodell mit zur Erforschung der Integration von Informationen aus multiplen Texten in der Lehrer:innenbildung und deren Zusammenhang mit epistemischen Variablen genutzt haben. Kindlinger (2021) argumentiert für die Notwendigkeit eines meta-reflexiven Ansatzes in der Lehrer:innenbildung und identifiziert Aspekte der Meta-Reflexivität in etablierten

Modellen zur Professionalisierung von Lehrer:innen, um die verschiedenen Ansätze konzeptuell zu integrieren. Dies stellt eine Verbindung zwischen meta-reflexiven Prinzipien und epistemologischen Überzeugungen her – das Desiderat einer breiten Operationalisierung von Meta-Reflexivität existiert jedoch weiterhin. Dieses noch ausstehende Desiderat greift der vorliegende Beitrag auf und dokumentiert die *Entwicklung einer Skala zur Erfassung der zugeschriebenen Bedeutsamkeit an meta-reflexive Prinzipien* für die Professionalisierung von (angehenden) Lehrpersonen. Der Beitrag macht sich dabei die Konkretion zunutze, die sich mit einer subjektiven Einschätzung der Bedeutsamkeit von Meta-Reflexivität für die individuelle Professionalisierung in der selbst erlebten Lehrer:innenbildung verbindet.

1. Einleitung

Der Anspruch, eine wissenschaftliche Theorie im Sinne eines komplexen sprachlichen Gebildes, „das in propositionaler oder begrifflicher Form die Phänomene eines Sachbereichs ordnet und die wesentlichen Eigenschaften der ihm zugehörigen Gegenstände und deren Beziehung untereinander“ (Thiel, 2018, S. 20) beschreibt, empirisch zu operationalisieren, ist ein komplexes Unterfangen. Es ist davon auszugehen, dass sich Theorie als ein solches konsistentes (allgemeines) Aussagesystem grundsätzlich in der Spezifik der Empirie bricht, da Handeln im sozialen Raum der Ungewissheit des Handlungsfeldes ausgesetzt ist: Was in der Theorie als allgemein ‚gültig‘ gesetzt wird, muss sich situativ in der Praxis nicht (zwingend) einstellen. Insofern kann eine Theorie auch nur bedingt empirisch ‚überprüft‘ werden bzw. ihre Relevanz oder Qualität hängt nicht generell von ihrer empirischen Erweisbarkeit ab. Der Beitrag beansprucht daher nicht die ‚Überprüfung‘ eines Theorie-Konstrukts *Meta-Reflexivität*, sondern fragt vielmehr danach, wie ausgehend vom theoretischen Postulat der Meta-Reflexivität auch empirisch geforscht werden kann, wobei die empirische Forschung möglichst theoriekonform erfolgen können soll. Dazu wird der Weg von der Theorie zu einem empirischen Instrument schrittweise nachgezeichnet.

Zunächst müssen die theoretischen Aussagen der Meta-Reflexivität in empirische Aussagen übertragen werden, zu welchen sich Befragte verhalten können. Denn in aller Regel treffen Theorien keine Aussagen über die Art und Weise, wie eine solche empirische Operationalisierung erfolgen kann. So überrascht es kaum, wenn zwei Forschende dieselbe Theorie in unterschiedlicher Weise und mit verschiedenen Zielsetzungen operationalisieren. Oder anders formuliert: „Die Theorie [...] zwingt [...] zu Festlegungen, die bis zu einem gewissen Grade beliebig sind“ (Eid et al., 2017, S. 41), wenn diese Theorie der empirischen Forschung zugänglich gemacht werden soll. Nach der Übertragung der Theoreme in ein empirisches Relativ stellt sich zudem die Frage, wie in

einem bestimmten Forschungsparadigma schließlich zu Daten gelangt wird. Im vorliegend gewählten quantitativ-empirischen Paradigma ist also zu klären, wie das empirische Relativ in ein numerisches Relativ überführt werden kann. Sollen die theoretischen Überlegungen der Meta-Reflexivität empirisch möglichst theoriekonform anschlussfähig gemacht werden, bedarf es einer homomorphen, d. h. strukturerhaltenden Abbildung zwischen den empirischen Aussagen und ihren numerischen Entsprechungen (vgl. z. B. Bortz & Döring, 2006). Werden diese Ausführungen auf die spezifische, hier vorgestellte empirische Operationalisierung von Meta-Reflexivität übertragen, ist hervorzuheben, dass es sich um *eine mögliche* Operationalisierung der empirisch niemals vollständig abbildbaren Theorie der Meta-Reflexivität handelt. Vorliegend wird also ein bestimmter Ausschnitt der Meta-Reflexivität operationalisiert – und nicht *das eine* theoretische Konstrukt Meta-Reflexivität (vgl. zu dieser Problematik Merk & Schmidt, 2023) – nämlich die Bedeutsamkeitszuschreibung von (angehenden) Lehrpersonen an die Prinzipien der Meta-Reflexivität für die Professionalisierung in der Lehrer:innenbildung.

Durch die nachfolgende Offenlegung der getroffenen Festlegungen im Prozess der Operationalisierung kann dieses Vorgehen selbst als meta-reflexiv angesehen werden, indem beispielsweise die Grenzen des quantitativ-empirischen Paradigmas mit reflektiert und damit transparent werden – womit explizit auch alternative Operationalisierungen im Blick sind (vgl. etwa die oben beschriebenen Studien: Cramer et al., 2023; Hartmann et al., 2021; Kindlinger, 2021). Der vorliegend gelegte Fokus auf die subjektive Bedeutsamkeitseinschätzung von meta-reflexiven Prinzipien für den Prozess der Professionalisierung aus Sicht von (angehenden) Lehrpersonen leitet sich aus der bereits theoretisch aufgeworfenen Annahme ab, dass eine Professionstheorie immer auch das berufliche Handeln selbst betrachten muss, womit eine empirische Perspektive triftig erscheint (Cramer et al., 2019). Die Operationalisierung leistet einen Beitrag dazu, die theoretische Perspektive der Meta-Reflexivität empirisch anschlussfähig zu machen und liefert damit Anhaltspunkte für ihre Validität und Hinweise auf die Akzeptanz aus Sicht der an der Lehrer:innenbildung Beteiligten. Denn eine Professionstheorie beansprucht einen ‚Sitz im Handlungsfeld‘ der Professionellen. Zudem ist eine Operationalisierung zwingend, um – ähnlich wie in Bezug auf viele andere Professionstheorien (vgl. Helsper, 2021) – eine empirische Forschung zur Meta-Reflexivität allererst zu ermöglichen. Vorliegend werden dabei Setzungen vorgenommen, die nicht alternativlos sind und daher zu einer hohen Transparenz verpflichtet: Es gibt aus meta-reflexiver Perspektive *a priori* keine Bevorzugung eines bestimmten empirischen Zugriffes auf die Theorie. Im Prozess der Operationalisierung müssen jedoch begründete Annahmen und Entscheidungen getroffen werden, ohne dass hieraus das Primat eines bestimmten Paradigmas zur empirischen Durchdringung von Meta-Reflexivität abzuleiten wäre.

Der Weg der Transformation meta-reflexiver Prinzipien (im Sinne postulierter Theoreme) in Items einer Skala zur Meta-Reflexivität (als Grundlage *einer möglichen* empirischen Erfassung) im Sinne deren Operationalisierung ist begründungsbedürftig und zentral für die inhaltliche Validität dieser Skala, die das zugrundeliegende theoretische Konstrukt der individuellen Bedeutsamkeitszuschreibungen an meta-reflexive Prinzipien in der Lehrer:innenbildung messen soll. Beim Bemühen, statistische Messmodelle im Hinblick auf bestimmte Kennwerte zu optimieren, muss der semantische Inhalt der zu messenden Konstrukte stets im Blick bleiben (vgl. Rossiter, 2008). Die hier vorgestellte meta-reflexive Skalenentwicklung richtet sich daher dezidiert an der Theorie und den daraus abgeleiteten meta-reflexiven Prinzipien aus. Daher werden nun zunächst die zu operationalisierenden theoretischen Prinzipien knapp rekapituliert und dann der Fokus für die gewählte Form der Operationalisierung begründet.

Die Arbeiten zur Meta-Reflexivität weisen vielfältige Bezüge auf, etwa wenn diese professionstheoretische Perspektive unter Rückgriff auf westliche Ideengeschichte, reflexionsbezogene Theoreme, den Diskurs um ‚personal epistemology‘, Komplexitätstheorie, Demokratiebildung oder andere Professions-theorien begründet wird (vgl. Cramer et al., 2019, 2023). Theoretisch wird ein *Prozessmodell* der rekursiven Relation der Sphäre des professionellen Handelns in der Schule zur Sphäre der (universitären) Lehrer:innenbildung postuliert, welches bereits für empirische Arbeiten herangezogen wurde (vgl. Hartmann et al., 2021; Kindlinger, 2021). Noch nicht gerieten bislang die vorliegend thematisierten, ebenfalls theoretisch postulierten *meta-reflexiven Prinzipien der Lehrer:innenbildung* hinsichtlich deren empirischen Operationalisierung in den Blick. Aktuell werden zehn Prinzipien der Meta-Reflexivität (für die Lehrer:innenbildung) postuliert (im Überblick: Cramer, 2020, S. 208). Zu den ursprünglich angenommenen (vgl. Cramer et al., 2019) zählen: (1) *Mehrperspektivität* auf Theorien und empirische Befunde; (2) *Distanzierung* als Erkennen des Mehrwerts distanzierter (zeitlich entlasteter) Betrachtungen; (3) *Kontextgebundenheit* wissenschaftlichen Erkennens in Theorie und Empirie; (4) *Alternativität* als Vorläufigkeit allen Erkennens; (5) *Unabhängigkeit* spezifischer Betrachtungsweisen, die in einem nicht hierarchischen Verhältnis zueinander stehen; (6) *Transparenz* als das Bewusstmachen der je spezifischen Axiomatik von Theorie und Empirie sowie des Modus meta-reflexiver Lehrer:innenbildung; (7) *Metakommunikation* über die Art und Weise, wie Theorie und Empirie meta-reflexiv betrachtet werden (können); (8) *Begründungspflicht* hinsichtlich der Auswahlkriterien und Quellen sowie Relevanz der herangezogenen Theorien und empirischen Befunde in der Lehrer:innenbildung; (9) *Dynamik* als schrittweise Erhöhung des Grades an Komplexität im Studienverlauf. Bereits dem Prinzip der Alternativität folgend wurde postuliert, dass künftig weitere Perspektiven hinzutreten können; eine solche wurde von Cramer et al. (2023) vorgeschlagen:

(10) *Historisierung* als Förderung der Sensibilität für die (kontingenten) historischen Umstände bei der Entstehung von Theorien und empirischen Befunden.

Initial diskutierte die Autor:innengruppe gemeinsam mit weiteren Expert:innen aus dem Bereich der Meta-Reflexivität und Skalenentwicklung verschiedene mögliche Wege, die meta-reflexiven Prinzipien empirisch zu operationalisieren. So könnte etwa die Realisierung dieser Prinzipien im Rahmen von Lehrveranstaltungen des Lehramtsstudiums als Zielgröße evaluativ erhoben werden. Alternative Überlegungen fokussierten auf Überzeugungen zur Relevanz der Umsetzung dieser Prinzipien in der Lehrer:innenbildung oder auf eine mögliche handlungsrelevante Funktion der Prinzipien im professionellen Lehrer:innenhandeln. Weiterhin wurden verschiedene Zielgruppen einer entsprechenden Erhebung abgewogen, wie beispielsweise Lehramtsstudierende, Dozierende, Referendar:innen, Auszubildende in Praxisphasen, Fortbildner:innen oder Lehrpersonen in der Berufstätigkeitsphase. Schlussendlich fiel die Entscheidung, in einem ersten Schritt die Einschätzung der Bedeutsamkeit („Wie wichtig...“) der Prinzipien im bildungswissenschaftlichen Teil des Lehramtsstudiums aus Sicht der Studierenden zu fokussieren. Neben forschungsökonomischen Gründen der schnellen und wiederholten Zugänglichkeit von Stichproben war dafür insbesondere ausschlaggebend, dass eine Konfrontation mit wissenschaftlichen Theorien und empirischen Befunden zuvorderst in der ersten Phase der Lehrer:innenbildung erfolgen und im Anschluss im Berufsverlauf sukzessive abnehmen dürfte.

In der Absicht, die theoretisch postulierte Relevanz der meta-reflexiven Prinzipien für den Prozess der Professionalisierung in der Lehrer:innenbildung aufzugreifen, werden die Proband:innen (hier Lehramtsstudierende mit Blick auf das Lehramtsstudium, prospektiv auch Referendar:innen und Lehrer:innen bezüglich der zweiten und dritten Phase der Lehrer:innenbildung) nach ihrer *Bedeutsamkeitszuschreibung* an die meta-reflexiven Prinzipien gefragt. Diese theoriekonforme Operationalisierung unterstellt, dass Einstellungen, Überzeugungen usw. zu Professionalität – wie auch immer sie sich hier letztlich fassen lassen (vgl. Merk & Schmidt, 2023) – von hoher Relevanz für den Professionalisierungsprozess selbst sind. Auch dann, wenn Professionstheorie und Empirie zur Professionalität auseinanderdriften (vgl. Cramer et al., 2023), stellt sich die Frage nach dem Verhältnis solcher Einschätzungen zur Meta-Reflexivität zu bereits etablierten und möglicherweise verwandten Konstrukten, wie etwa epistemischen Überzeugungen, deren Relevanz für die Lehrer:innenbildung wiederholt gezeigt wurde (z. B. Merk, 2020; Rosman et al., 2020; Skott, 2015; Wilde & Kunter, 2016).

Der Beitrag folgt damit den *Fragestellungen*, wie die theoretisch postulierten Prinzipien von Meta-Reflexivität empirisch operationalisiert werden können, welche skalenanalytischen Ergebnisse ein solcher Versuch der Operationalisierung zeigt und in welchem Verhältnis die so gewonnenen Daten mit an-

grenzenden und empirisch etablierten Konstrukten (hier: epistemischen Überzeugungen) stehen. Zu deren Bearbeitung wird zunächst der Weg von den theoretisch postulierten Prinzipien der Meta-Reflexivität hin zu entsprechenden Itemformulierungen für ein Erhebungsinstrument nachgezeichnet. Davon ausgehend wird der Weg der Skalenentwicklung über vier Pilotierungsstudien hinweg aufgezeigt. Die so sukzessive gewonnene finale Skala wird anhand der sich empirisch zeigenden drei latenten Faktoren inhaltlich beschrieben und an die Theorie rückgebunden. Bezüge zu den zusätzlich erhobenen epistemischen Überzeugungen werden aufgezeigt. Die Ergebnisse werden auf Basis von ausgewiesenen Limitationen diskutiert und Perspektiven für weiterführende Forschung aufgezeigt.

2. Prozess der empirischen Operationalisierung meta-reflexiver Prinzipien

Die empirische Operationalisierung der meta-reflexiven Prinzipien erfolgt in vier diskreten Pilotierungsstudien, die nachfolgend im Detail dargestellt werden.

2.1 Pilotierung 1: Itementwicklung und konfirmatorische Analysen

Orientiert an den Gesichtspunkten für gute Itemformulierungen (Moosbrugger & Brandt, 2020, S. 75–78) wurde ein Pool von Items in Form von Aussagen zu den ursprünglich neun meta-reflexiven Prinzipien erstellt. Anschließend wurden vier Items pro Prinzip für die Pilotierung ausgewählt, um trotz einer potenziellen Kürzung anhand von Item-Trennschärfen mit einer ausreichenden Zahl an Items zu beginnen. Der Fragetext zur Einschätzung der Items lautete: „Für wie wichtig erachten Sie das Folgende in Ihrem Bildungswissenschaftlichen Studium?“. Für die Beantwortung stand eine 6-stufige Antwortskala mit beschrifteten Extrempolen (1 = „überhaupt nicht wichtig“ bis 6 = „sehr wichtig“) zur Verfügung. Tabelle 1 bietet nachfolgend eine Übersicht über die verwendeten Items.

Tabelle 1: Übersicht zu den in Pilotierung 1 eingesetzten Items

1MP	Die behandelten Inhalte werden aus mehreren Perspektiven betrachtet.
2MP	Verschiedene Blickwinkel auf einen bestimmten Gegenstand werden eingenommen.
3MP	Bei der Betrachtung eines Themas werden unterschiedliche Sichtweisen abgewogen.
4MP	Zur Auseinandersetzung mit den Fragestellungen werden vielfältige Perspektiven herangezogen.
5KG	Bei der Behandlung wissenschaftlicher Positionen wird deutlich, auf welchen Annahmen diese beruhen.

- 6KG Der Geltungsraum von bestimmten wissenschaftlichen Erkenntnissen wird thematisiert.
- 7KG Die Herkunft verschiedener wissenschaftlicher Denkweisen wird behandelt.
- 8KG Der Weg des Zustandekommens bestimmter Annahmen wird erläutert.
- 9AL Erkenntnisse zu einem Gegenstand werden als vorläufig gekennzeichnet.
- 10AL Die Art und Weise der Beschäftigung mit einem Thema wird als eine Möglichkeit neben anderen angesehen.
- 11AL Bei der Arbeit an Inhalten wird die Vorläufigkeit der getroffenen Schlussfolgerungen thematisiert.
- 12AL Die getroffenen Aussagen werden als nur eine Möglichkeit angesehen, sich zu den Inhalten zu äußern.
- 13UA Unterschiedliche Ansichten müssen nicht zwingend im Widerspruch zueinander stehen.
- 14UA Verschiedene Sichtweisen auf einen Sachverhalt werden unvoreingenommen nebeneinandergestellt.
- 15UA Eine bestimmte Perspektive auf einen Sachverhalt wird nie ausschließlich als ‚richtig‘ oder ‚falsch‘ bezeichnet.
- 16UA Allein einen ganz bestimmten Blickwinkel auf ein Thema einzunehmen, wird als unzureichend thematisiert.
- 17MK Es wird kommuniziert, wie mit verschiedenen Perspektiven auf einen Gegenstand umgegangen wird.
- 18MK Der Umgang mit unterschiedlichen Positionen zu Themen wird besprochen.
- 19MK Es wird die Art und Weise thematisiert, wie mit unterschiedlichen Sichtweisen umgegangen werden kann.
- 20MK Wie verschiedene Perspektiven auf einen Gegenstand behandelt werden, wird thematisiert.
- 21DI Die behandelten Inhalte werden aus der Distanz heraus betrachtet.
- 22DI Eine skeptische Grundhaltung gegenüber den thematisierten Inhalten ist gegeben.
- 23DI Die Auseinandersetzung mit den Themen erfolgt distanziert und unvoreingenommen.
- 24DI Es erfolgt eine kritische Prüfung der behandelten Inhalte.
- 25TR Das Vorgehen zum Umgang mit verschiedenen Sichtweisen wird offengelegt.
- 26TR Die Art und Weise der Auseinandersetzung mit unterschiedlichen Perspektiven wird thematisiert.
- 27TR Formen der Beschäftigung mit alternativen Sichtweisen auf einen Sachverhalt werden angesprochen.
- 28TR Die Vorgehensweise beim Auftreten von alternativen wissenschaftlichen Einschätzungen wird besprochen.
- 29BP Es wird begründet, warum bestimmte Inhalte und Quellen herangezogen werden.
- 30BP Die Auswahlkriterien für die behandelten Gegenstände und Materialien werden offengelegt.
- 31BP Es wird erläutert, warum die jeweiligen Themen wichtig sind.
- 32BP Es wird besprochen, inwiefern bestimmte Texte zur Bearbeitung eines Themas geeignet sind.
- 33DY Es wird darauf geachtet, dass die Komplexität der behandelten Sichtweisen nach und nach zunimmt.

- 34DY Die Vielfalt der thematisierten Perspektiven auf einen Gegenstand wird schrittweise gesteigert.
- 35DY Die Zusammenhänge unterschiedlicher Betrachtungsweisen werden stufenweise mehr verdeutlicht.
- 36DY Die Komplexität der behandelten Fragestellungen tritt nach und nach hervor.

Skalen: MP = Mehrperspektivität; KG = Kontextgebundenheit; AL = Alternativität; UA = Unabhängigkeit; MK = Metakommunikation; DI = Distanzierung; TR = Transparenz; BP = Begründungspflicht; DY = Dynamik.

Die Erhebung wurde als Online-Fragebogen konzipiert, bei dem zur Vermeidung möglicher Reihenfolgeeffekte bei der Itembeantwortung (Moosbrugger & Brandt, 2020, S. 85 f.) eine randomisierte Abfolge der Items spezifiziert wurde. Für die erste Pilotierung wurden insgesamt $N_1 = 167$ Lehramtsstudierende an der Universität Tübingen sowie an der Pädagogischen Hochschule Vorarlberg befragt, die sich zumeist in der Anfangsphase ihres Studiums befanden (mittlere Studiendauer: 2.8 Semester). Die Befragung erfolgte im Wintersemester 2021/2022 sowie Sommersemester 2022 in Lehrveranstaltungen, wenige vor Ort fehlende Studierende füllten den Fragebogen zeitverzögert aus. Zur Konstruktvalidierung der entwickelten Skalen wurden Konfirmatorische Faktorenanalysen mithilfe des Packages *lavaan* (Rosseel, 2012) in der Statistiksoftware R durchgeführt. Die Modellschätzung erfolgte mithilfe des Maximum-Likelihood-Schätzers mit robusten Standardfehlern und Yuan-Bentler-Korrektur für die Teststatistiken (Robitzsch, 2020).

Da aufgrund theoretischer Vorüberlegungen von hohen Korrelationen zwischen den Items zu den unterschiedlichen Prinzipien ausgegangen wurde, erfolgte in einem ersten Schritt die separierte Testung der einzelnen Subskalen, welche die theoretisch angenommenen Prinzipien mit jeweils vier Items repräsentieren, hinsichtlich deren Eindimensionalität. Ausgehend von den Beurteilungskriterien von Gäde et al. (2020) weisen die Ergebnisse durchgehend auf eine akzeptable bis gute interne Konsistenz und eine hohe Modellgüte für sieben der neun Subskalen hin (Details online: <https://doi.org/10.17605/OSF.IO/BFRG5>). Die Subskala für das Prinzip der Alternativität zeigt bei ansonsten akzeptablen Fit-Statistiken einen etwas erhöhten RMSEA-Wert von .099. Trotz einer akzeptablen Reliabilität liegen die Passungswerte für die Subskala der Distanzierung durchgehend außerhalb des akzeptablen Bereichs, wobei sich in vertiefenden Analysen hohe Residualkorrelationen zwischen drei der vier Items zeigen. Die strukturelle Testung des Gesamtmodells mit den neun Subskalen weist aufgrund eines nicht akzeptablen CFI-Werts von 0.883 auf eine schlechte Passung der angenommenen Gesamtstruktur zu den Daten hin (vgl. Details online). Weiterhin ergeben sich die erwarteten hohen Korrelationen ($.655 < r < .956$) zwischen den Subskalen (vgl. Details online), sodass alternativ ein einfaktorielles Modell getestet wurde, bei dem alle 36 Items des Instruments

auf einen latenten Faktor laden. Dieses Modell zeigte eine signifikant schlechtere Passung als das 9-faktorielle Gesamtmodell ($\Delta X^2(36) = 140.6, p < .001$).

Die hohen Interkorrelationen zwischen den Subskalen lassen sich nicht nur durch inhaltlich-konstruktbezogene Überschneidungen erklären, sondern begründen sich teilweise auch in der hohen sprachlich-semantischen Ähnlichkeit zwischen Items unterschiedlicher Prinzipien. Beispielsweise findet sich der Begriff der *Sichtweise* in insgesamt sechs der 36 Items über die Subskalen Mehrperspektivität, Unabhängigkeit, Metakommunikation, Transparenz und Dynamik hinweg. Ebenso wird der semantisch ähnliche Begriff *Perspektive* insgesamt siebenmal verwendet. Auf Grundlage dieser Ergebnisse wurde für das weitere Vorgehen die Annahme einer neun-faktoriellen Struktur datengetrieben aufgegeben und eine Kurzsкала mit den neun trennschärfsten Items entwickelt. Dieses Instrument wies auf dieser Datengrundlage eine gute interne Konsistenz auf ($\alpha = .866; \omega = .865$).

2.2 Pilotierung 2: Kurzsкала zur Testung einer einfaktoriellen Lösung

Eine eindimensionale Lösung, bei der die neun theoretisch postulierten Prinzipien das zu messende Konstrukt gemeinsam und gleichwertig repräsentieren, erscheint theoretisch durchaus plausibel. Aufgrund der guten internen Konsistenz der Kurzsкала mit den neun trennschärfsten Items aus der ersten Pilotierung sollte daher in einer zweiten Pilotierung überprüft werden, ob diese Kurzsкала eine gemeinsame latente Variable adäquat misst. Für jedes der neun Prinzipien wurde eine Subskala aus den vier im Fragebogen der ersten Pilotierung enthaltenen Items gebildet, dann wurden mittels *lavaan* in R modellbasierte Reliabilitätsschätzungen vorgenommen (Schermelleh-Engel & Gädde, 2020). Das jeweils trennschärfste Item repräsentiert aus testtheoretischer Perspektive die jeweils zugrundeliegende Subskala (semantisch: das meta-reflexive Prinzip) angemessen. Yousfi (2005) weist auf mögliche Schwierigkeiten dieses Vorgehens der klassischen Testtheorie hin. Hohe interne Konsistenzen der Kurzsкала werden daher als notwendige, nicht jedoch bereits hinreichende Voraussetzung für das Vorliegen einer eindimensionalen Lösung betrachtet. Bei der Erprobung dieses Kurzinstrumentes (ohne die Erfassung der weiteren Items aus dem Gesamtpool von Pilotierung 1) mittels einer Befragung von $N_2 = 164$ Lehramtsstudierenden der Universität Tübingen (Schnittmenge zur initialen Befragung aus Pilotierung 1) in der Studieneingangsphase ergibt sich jedoch eine weitaus schlechtere, nicht mehr akzeptable interne Konsistenz ($\alpha = .625; \omega = .606$). Damit wird eine eindimensionale Kurzsкала mit neun Items im weiteren Vorgehen verworfen.

2.3 Pilotierung 3: Neuausrichtung und explorative Faktorenanalyse

Die meta-reflexiven Prinzipien lassen sich empirisch weder alle einem gemeinsamen Faktor zuordnen, noch zerfällt das hier untersuchte meta-reflexive Konstrukt in neun trennscharfe Dimensionen, die den theoretisch postulierten Dimensionen entsprechen. Diese Erkenntnisse implizieren eine Loslösung von der Annahme der statistischen Trennbarkeit aller theoretisch postulierten Prinzipien mit Blick auf die hier in Rede stehende Bedeutsamkeitszuschreibung. Stattdessen wird zur datengestützten Identifizierung latenter Faktoren sowie deren inhaltlichen Beschreibung übergegangen. Anstelle wie bisher strukturprüfend, wird nun struktursuchend mittels explorativer Faktorenanalyse (EFA) vorgegangen. In einer Abwägung zwischen Testökonomie (höher bei weniger Items) und (erwartbarer) Reliabilität (höher bei mehr Items), werden zwei Items pro Prinzip herangezogen, indem zu den bereits bestehenden Items der zweiten Pilotierung noch die jeweils zweit-trennschärfsten Items aus der ersten Pilotierung ergänzt werden. Zwischenzeitlich trat im Prozess der Theoriebildung das weitere meta-reflexive Prinzip der *Historisierung* hinzu (vgl. Abschnitt 1), welches in dieser Pilotierung zusätzlich berücksichtigt wird, wodurch die neue Skala 20 Items – zehn theoretisch postulierte Prinzipien mit jeweils zwei Items – umfasst. So soll die theoretisch-inhaltliche Validität der Skala auch durch eine konsequent an den theoretischen Prinzipien ausgerichtete Itementwicklung sichergestellt werden. Als theoretisch begründete normative Setzung verbleibt als Kriterium zur Beurteilung der Modellpassung, dass die jeweilige Primärladung der beiden Items desselben theoretisch postulierten meta-reflexiven Prinzips nicht auf unterschiedlichen Faktoren liegen soll. Items werden im Verlauf der EFA ausgeschlossen, wenn die Ladung auf keinem der Faktoren ausreichend hoch ist. Als Cut-off-Wert wird gemäß der Literatur ein Absolutbetrag von $\lambda = .30$ herangezogen (z. B. Cudek, 2000).

Die Pilotierung 3 fand im vorliegenden Kontext im Wintersemester 2022/2023 an der Universität Tübingen sowie an der Pädagogischen Hochschule Freiburg statt und die Stichprobe umfasste $N_3 = 139$ Lehramtsstudierende in der Studienanfangsphase. Die Arbeitsschritte der durchgeführten EFA richten sich methodisch nach Brandt (2020), die Berechnungen erfolgten mithilfe von diversen R-Packages (Jorgensen et al., 2022; Korkmaz et al., 2014; Revelle, 2023; Rosseel, 2012). Zur besseren intersubjektiven Überprüfbarkeit sind die R-Markdowns für diese und die nachfolgende Pilotierung online zugänglich: <https://doi.org/10.17605/OSF.IO/BFRG5>.

Zur Bestimmung der Faktorenanzahl können unterschiedliche Ansätze in einer EFA verwendet werden. Häufig eingesetzt wird die Maximum-Likelihood-Faktorenanalyse, da diese – ähnlich wie bei konfirmatorischen Analysen – eine inferenzstatistische Überprüfung eines Modells erlaubt. Eine Vorausset-

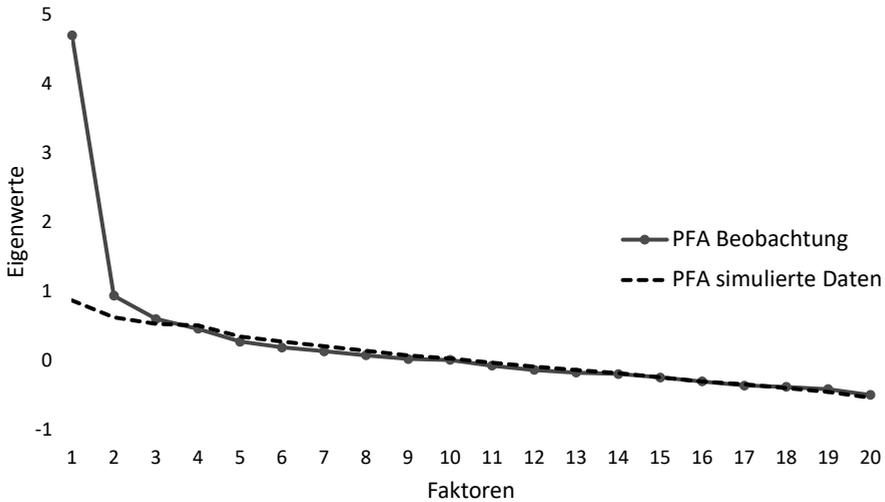


Abbildung 1: Parallelanalyse zu Pilotierung 3

zung zur Anwendung dieser Methode ist die multivariate Normalverteilung der Daten (Eid et al., 2017, S. 899). Wie bei den vorherigen Pilotierungen fällt deskriptiv eine im Mittel hohe Zustimmung zu den Items in der untersuchten Stichprobe auf. Jedes Item wurde daraufhin inferenzstatistisch auf Vorliegen eines normalverteilten Antwortverhaltens untersucht. Bei einigen Items zeigte sich eine Verletzung der Normalverteilungsannahme und auch die drei im R-Package *MVN* enthaltenen Tests auf multivariate Normalverteilung (Korkmaz et al., 2014) kamen alle zu dem Ergebnis, dass die Daten nicht multivariat normalverteilt sind. Daher wird vorliegend eine Hauptachsenanalyse (*Principal Factor Analysis*, PFA) durchgeführt. Nach dem Kaiser-Kriterium legen sich fünf Faktoren als relevant nahe, da deren Eigenwerte größer als eins sind. Dieses Kriterium neigt aber zu einer Überfaktorisierung (vgl. Fabrigar et al., 1999), weshalb die fünf extrahierten Faktoren als obere Schranke der Anzahl relevanter Faktoren zu verstehen ist. Im Scree-Plot ist ein deutlicher Knick nach dem ersten Faktor zu sehen, weshalb nach dem Scree-Kriterium nach Cattell (1966) eine einfaktorische Lösung plausibel erscheint. Die Parallelanalyse nach Horn (1965) wird in der Literatur als das zuverlässigste Kriterium der Faktorextraktion angesehen (z. B. Dinno, 2009; Timmerman & Lorenzo-Seva, 2011), da hier jene Faktoren als bedeutsam ausgewählt werden, deren Eigenwerte größer sind als die Eigenwerte zufällig generierter Daten, die sich lediglich in der Eigenschaft der Unkorreliertheit aller Variablen vom empirischen Datensatz unterscheiden. Je nach Simulation weist die Parallelanalyse (eine exemplarische Durchführung ist in Abb. 1 dargestellt) auf eine zwei- oder dreifaktorielle Lösung hin.

Eine Unkorreliertheit der Faktoren ist weder theoretisch plausibel noch durch die bisherigen Pilotierungen empirisch anzunehmen, sodass eine oblique Faktorenrotation adäquater ist als eine orthogonale Rotation (Fabrigar et al., 1999). Mittels Oblimin-Rotation wird versucht, eine Einfachstruktur der Faktorladungsmatrix (Brandt, 2020, S. 596–597) herzustellen, die eine möglichst eindeutige Zuordnung zwischen Items und Faktoren und damit eine gute inhaltliche Interpretierbarkeit der Faktoren ermöglicht. Die zweifaktorielle Lösung ergibt ein diffuses Bild, da mehrfach Items desselben meta-reflexiven Prinzips auf unterschiedlichen Faktoren laden. Im Falle der dreifaktoriellen Lösung werden zunächst zwei Items ausgeschlossen, die auf keinem der drei Faktoren ausreichend hoch laden. Aus inhaltlichen Gründen werden vier weitere Items zu vier unterschiedlichen Prinzipien ausgeschlossen, weil diese uneindeutige Mehrfachladungen auf mindestens zwei Faktoren aufweisen. Da das gemessen an den theoretischen Vorannahmen korrespondierende zweite Item der betreffenden Prinzipien jeweils eindeutig auf einem Faktor lädt, bleiben alle theoretisch postulierten Prinzipien weiterhin in der Skala repräsentiert. Eine Rotation ohne die sechs ausgeschlossenen Items führt zu einem klareren Bild: Lediglich ein Item lädt nun nicht mehr ausreichend hoch auf einem der drei Faktoren und wird daher ebenfalls ausgeschlossen. In Tabelle 2 sind die 13 verbliebenen Items dargestellt, die sich auf drei empirisch identifizierte Faktoren verteilen.

Eine inferenzstatistische Schätzung zum Modellfit kann bei der PFA nicht vorgenommen werden, sodass auch die Indices einer konfirmatorischen Faktorenanalyse nicht sinnvoll interpretiert werden könnten. Es bleibt der Blick auf die Residualmatrix als Differenz zwischen der empirischen und modellimplizierten Korrelationsmatrix. Diese weist Einträge betragsmäßig kleiner .10 auf, was auf einen guten Modellfit hindeutet (Brandt, 2020, S. 604). Für die drei identifizierten Faktoren wurde jeweils McDonalds Omega modellbasiert geschätzt (Schermelleh-Engel & Gåde, 2020): $\omega_1 = .740$, $\omega_2 = .658$ bzw. $\omega_3 = .630$. Mithilfe der Spearman-Brown-Formel kann die für eine gewünschte Reliabilität erforderliche Anzahl an Testitems berechnet werden. Bei einer angestrebten Anhebung der Omega-Werte auf etwa .800 sind zehn, neun bzw. fünf Items zur Messung der Faktoren notwendig, die von der derzeitigen Skala (vgl. Tab. 2) noch unterlaufen werden, was eine Ergänzung von Items zu den drei Subskalen im nächsten Schritt erfordert.

Tabelle 2: Oblimin-rotierte Faktorladungsmatrix der Items in Pilotierung 3

Item (theoretisches Prinzip)	Faktor 1 <i>Vorläufigkeit</i>	Faktor 2 <i>Multi-perspektivität</i>	Faktor 3 <i>Erkenntnis-genese</i>
Bei der Arbeit an Inhalten wird die Vorläufigkeit der getroffenen Schlussfolgerungen thematisiert. (<i>Alternativität</i>)	.623		
Die Vorgehensweise beim Auftreten von alternativen wissenschaftlichen Einschätzungen wird besprochen. (<i>Transparenz</i>)	.519		
Der Umgang mit unterschiedlichen Positionen zu Themen wird besprochen. (<i>Metakommunikation</i>)	.326		
Die Auswahlkriterien für die behandelten Gegenstände und Materialien werden offengelegt. (<i>Begründungspflicht</i>)	.565		
Es wird begründet, warum bestimmte Inhalte und Quellen herangezogen werden. (<i>Begründungspflicht</i>)	.440		
Die Vielfalt der thematisierten Perspektiven auf einen Gegenstand wird schrittweise gesteigert. (<i>Dynamik</i>)	.654		
Die Komplexität der behandelten Fragestellungen tritt nach und nach hervor. (<i>Dynamik</i>)	.438		
Verschiedene Blickwinkel auf einen bestimmten Gegenstand werden eingenommen. (<i>Mehrperspektivität</i>)		.515	
Zur Beschäftigung mit den Fragestellungen werden vielfältige Perspektiven herangezogen. (<i>Mehrperspektivität</i>)		.880	
Eine kritische Grundhaltung gegenüber den thematisierten Inhalten ist gegeben. (<i>Distanzierung</i>)		.412	
Allein einen ganz bestimmten Blickwinkel auf ein Thema einzunehmen, wird als unzureichend thematisiert. (<i>Unabhängigkeit</i>)		.322	
Der historische Entstehungsprozess von Theorien oder empirischen Befunden wird angesprochen. (<i>Historisierung</i>)			.473
Die Herkunft verschiedener wissenschaftlicher Denkweisen wird behandelt. (<i>Kontextgebundenheit</i>)			.853

Stichprobe: $N_3 = 139$. *Frage:* Für wie wichtig erachten Sie das Folgende in Ihrem Bildungswissenschaftlichen Studium? *Antwortformat:* 6-stufige Likert-Skala mit beschrifteten Extrempolen (1 = „überhaupt nicht wichtig“; 6 = „sehr wichtig“).

2.4 Pilotierung 4: Bestätigung der Ergebnisse und vorläufiges Modell

Im iterativen Operationalisierungsprozess wurde in der Absicht einer reliablen Erfassung der drei datengetriebenen identifizierten Faktoren, die empirischen Prinzipien von Meta-Reflexivität entsprechen, erneut eine Itementwicklung erforderlich (vgl. 2.3). Diese basiert nun auf den drei empirisch identifizierten Faktoren und nicht mehr auf den theoretisch postulierten Prinzipien von Meta-Reflexivität. Die in Tabelle 2 dargestellten 13 Items wurden wortgleich übernommen, dazu wurden drei (Faktor 1), fünf (Faktor 2) bzw. drei (Faktor 3) weitere Items ergänzt (teils aus dem anfänglichen Pool der ersten Pilotierung, teils infolge Neuformulierung korrespondierend zu den verbliebenen Items in Pilotierung 3), sodass die neue Skala insgesamt 24 Items umfasst. Bis auf die neu hinzugefügten Items war der in einer neuen, unabhängigen Stichprobe eingesetzte Fragebogen inhaltsgleich wie bei der dritten Pilotierung. Die Befragung fand im Mai 2023 erneut unter Studienanfänger:innen des Lehramts an der Universität Tübingen statt ($N_4 = 99$). Im Wesentlichen konnten die Befunde der vorherigen Pilotierung repliziert werden, wie dem ausführlichen R-Markdown (<https://doi.org/10.17605/OSF.IO/BFRG5>) entnommen werden kann. Zunächst werden vier Items ausgeschlossen, da diese nicht ausreichend hoch auf dem Faktor laden, dem sie theoretisch zugeordnet sind, und weil sie zudem auf einem der beiden anderen Faktoren relativ stark laden. Danach wurde ein weiteres Item aus der Analyse genommen, da dieses nun eindeutig auf einem der beiden anderen Faktoren lädt. Ein Item verbleibt aufgrund der Ergebnisse in Pilotierung 3 in der Skala, obwohl es in dieser untersuchten Stichprobe nicht mehr nur auf einem Faktor lädt, ein weiteres Item wird aufgrund einer solchen Doppelladung ausgeschlossen, da sich die interne Konsistenz der betrachteten Subskala nach Entfernung dieses Items erhöht.

Nach Ausschluss von insgesamt sechs Items zeigte sich eine dreifaktorielle Lösung erneut als empirisch stabil. Die so gewonnene *Skala zur Erfassung der Bedeutsamkeitszuschreibung an meta-reflexive Prinzipien in der Lehrer:innenbildung* umfasst damit jeweils sieben Items für die ersten beiden sowie vier Items für den dritten Faktor, insgesamt also 18 Items (vgl. Tab. 3).

Etwa 34 Prozent der aufgetretenen Varianz kann durch die drei Faktoren erklärt werden. In Bezug auf die angestrebte Steigerung der internen Konsistenzen zeigte sich insgesamt betrachtet eine signifikante Verbesserung: Für den ersten Faktor sank die interne Konsistenz nur geringfügig auf $\omega_1 = .729$ (zuvor $.740$), die größere Itemanzahl der beiden anderen Faktoren führte jedoch zu entsprechend höheren Omega-Werten von $\omega_2 = .769$ (vorher $.658$) bzw. $\omega_3 = .697$ (zuvor $.630$). Die nun vorliegende Skala muss zwar insgesamt als vorläufig erachtet werden, ist aber tragfähig und kann daher in diesem Beitrag als finale Skala angesehen werden, mit der künftig geforscht werden kann.

Tabelle 3: Oblimin-rotierte Faktorladungsmatrix der Items in Pilotierung 4 (finale Skala)

Itemtext	Faktor 1 Vorläufig- keit	Faktor 2 Multi- perspektivität	Faktor 3 Erkenntnis- genese
Bei der Arbeit an Inhalten wird die Vorläufigkeit der getroffenen Schlussfolgerungen thematisiert.	.492		
Die Vorgehensweise beim Auftreten von alternativen wissenschaftlichen Einschätzungen wird besprochen.	.608		
Der Umgang mit unterschiedlichen Positionen zu Themen wird besprochen.	.289		
Die Auswahlkriterien für die behandelten Gegenstände und Materialien werden offengelegt.	.224		
Es wird begründet, warum bestimmte Inhalte und Quellen herangezogen werden.	.506		
Die Vielfalt der thematisierten Perspektiven auf einen Gegenstand wird schrittweise gesteigert.	.642		
Die Komplexität der behandelten Fragestellungen tritt nach und nach hervor.	.561		
Verschiedene Blickwinkel auf einen bestimmten Gegenstand werden eingenommen.		.651	
Eine kritische Grundhaltung gegenüber den thematisierten Inhalten ist gegeben.		.466	
Allein einen ganz bestimmten Blickwinkel auf ein Thema einzunehmen, wird als unzureichend thematisiert.		.348	
Die behandelten Inhalte werden aus mehreren Perspektiven betrachtet.*		.669	
Bei der Betrachtung eines Themas werden unterschiedliche Sichtweisen abgewogen.*		.590	
Bei der Betrachtung der vorgestellten Inhalte wird eine kritische Distanz eingenommen.*		.621	
Es wird eine kritische Prüfung der vorgestellten Theorien und empirischen Befunde vorgenommen.*		.496	
Der historische Entstehungsprozess von Theorien oder empirischen Befunden wird angesprochen.			.538
Die Herkunft verschiedener wissenschaftlicher Denkweisen wird behandelt.			.388
Wissenschaftliche Erkenntnisse werden in ihrem jeweiligen Zeitgeist verortet.*			.638
Es wird auf die jeweiligen zeitgeschichtlichen Umstände bei der Entstehung wissenschaftlicher Erkenntnisse aufmerksam gemacht.*			.710

Anmerkung: *in Pilotierung 4 hinzugenommene Items. Stichprobe: $N_1 = 99$. Frage: Für wie wichtig erachten Sie das Folgende in Ihrem Bildungswissenschaftlichen Studium? Antwortformat: 6-stufige Likert-Skala mit beschrifteten Extrempolen (1 = „überhaupt nicht wichtig“; 6 = „sehr wichtig“).

2.5 Konvergente Validierung: Meta-Reflexivität und Epistemische Überzeugungen

Die gewählte Form der Operationalisierung meta-reflexiver Prinzipien muss über die faktorenanalytische Skalenentwicklung (vgl. 2.1 bis 2.4) hinaus auch konstruktvalide sein, was empirisch durch das Vorliegen konvergenter und diskriminanter Validität nachgewiesen werden kann (vgl. Eid et al., 2017, S. 75). Für den Nachweis der konvergenten Validität wird idealtypisch dasselbe Konstrukt über verschiedene Erhebungsmethoden gemessen, und die verschiedenen Ergebnisse korrelieren hoch positiv miteinander. Daher wurde die dritte Pilotierung um eine Skala aus einem theoretisch nahen Forschungsfeld ergänzt, die im Idealfall ein Konstrukt misst, das auch durch die eigene Skala zur Einschätzung der meta-reflexiven Prinzipien mit erfasst wird. Dies trifft auf das Forschungsfeld der *Personal Epistemology* zu, da in beiden Fällen Kognitionen über epistemische Entitäten thematisiert werden. Meta-Reflexivität versucht aus professionstheoretischer Perspektive, Zielkategorien epistemischer Überzeugungen zu formulieren, die für professionelles Handeln relevant sind. Ansätze der ‚Personal Epistemology‘ hingegen beschreiben epistemische Überzeugungen empirisch und enthalten damit auch nichtwissenschaftstheoretische, gewissermaßen naive Überzeugungen (vgl. Cramer et al., 2019, S. 413f.). Theoretisch sind somit Überschneidungen der beiden Konstrukte vorhanden, die sich auch empirisch zeigen könnten.

Krettenauer (2005) postuliert drei Hauptstufen einer intraindividuellen Entwicklung epistemischer Überzeugungen und verwendet hierfür die Bezeichnungen des Absolutismus, Relativismus und Postrelativismus. Eine *absolutistische* Überzeugung wäre, Wissen als eindeutig wahr oder falsch zu bezeichnen. Werden Meinungen oder Alltagstheorien zur Begründung von Wissen herangezogen, spricht man von *relativistischen* Überzeugungen. Der Übergang zu *postrelativistischen* Überzeugungen ist dadurch charakterisiert, dass Wissen weiterhin „[...] als nicht eindeutig wahr oder falsch angesehen wird, aber die Überzeugung vorliegt, dass eine Evaluation des Gültigkeitsanspruchs durch Untersuchung der Argumente und des Vergleichs von Evidenz möglich ist.“ (Merk et al., 2017, S. 172).

Dass eine Zustimmung zu postrelativistischen Auffassungen mit einer hohen Bedeutsamkeitseinschätzung bestimmter meta-reflexiver Prinzipien einhergeht, erscheint plausibel (vgl. Kindlinger, 2021, S. 44–45). Schon in Pilotierung 3 (vgl. 2.3) wurde die Erhebung zum Zwecke der Konstruktvalidierung um eine Variation der ursprünglich von Krettenauer (2005) entwickelten Skala zur Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen (FREE) ergänzt, nämlich um die Skala FREE-EDPSY (Rosman et al., 2019, 2020). In Pilotierung 4 konnten die zuvor in Pilotierung 3 bereits sichtbaren Zusam-

menhänge zwischen den Dimensionen beider Konstrukte weitgehend repliziert werden (vgl. Tab. 4).

Tabelle 4: Interkorrelationen zwischen den Faktoren meta-reflexiver- und epistemologischer Überzeugungen (Pilotierung 4)

	Vorläufig- keit	Multiperspektivität	Erkenntnis- genese
Multiperspektivität	.362**		
Erkenntnisgenese	.249**	.241**	
Absolutismus	.142	-.275**	-.025
Relativismus	.178	-.083*	.041
Postrelativismus	.184*	.443**	.294*

Anmerkungen: ** $p < .01$; * $p < .05$ (einseitig).

Die meta-reflexiven Faktoren untereinander korrelieren theoriekonform moderat positiv, ebenso schlüssig ist die durchweg positive Korrelation mit postrelativistischen Überzeugungen. Während die meta-reflexiven Faktoren Vorläufigkeit und Erkenntnisgenese nur schwach, aber erwartungskonform mit den epistemologischen Überzeugungen zusammenhängen, korreliert Multiperspektivität signifikant mit den im FREE-EDPSY beschriebenen Entwicklungsstufen epistemologischer Überzeugungen, und zwar stark positiv mit Postrelativismus und negativ mit Absolutismus. Studierende, die meta-reflexiven Prinzipien in der Lehrer:innenbildung höhere Bedeutung zuschreiben, können insgesamt als stärker postrelativistisch und weniger absolutistisch beschrieben werden.

3. Diskussion

In der ersten Pilotierung konnte die Annahme der Eindimensionalität für die meisten der neun Subskalen verifiziert werden, die je vier eingesetzten Items repräsentierten das zugrundeliegende theoretische Prinzip adäquat. Auch aufgrund der hohen Interkorrelationen zwischen den Prinzipien wurden die neun trennschärfsten Items (eines pro Prinzip) zu einer gemeinsamen Skala zusammengefasst. Bei Verwendung der Daten aus der ersten Pilotierung zeigte sich eine hohe interne Konsistenz, in der zweiten Pilotierung konnte dieser Befund nicht repliziert werden. Die theoretisch plausible einfaktorielle Lösung war damit bis auf Weiteres empirisch falsifiziert. Ab der dritten Pilotierung wurde daher konsequenterweise explorativ vorgegangen mit Interesse an der faktoriellen Struktur der Bedeutsamkeitseinschätzung der meta-reflexiven Prinzipien. Im Ergebnis wurde eine dreifaktorielle Lösung vorgeschlagen. Nach erneuter

Itementwicklung bestätigte die vierte Pilotierung dieses Ergebnis mit einer 18 Items umfassenden Skala.

Zur Benennung und Interpretation dieser drei empirisch identifizierten Faktoren lassen sich zweierlei Argumentationslinien anführen: Einerseits kann von der initialen meta-reflexiven Theorie her gedacht werden, andererseits bietet sich das Begriffsinventar der Forschung zu Personal Epistemology hierfür an:

Vorläufigkeit: Aus theoretischer Sicht der *Meta-Reflexivität* verbinden sich mit dem ersten Faktor (empirisch) bzw. der ersten Dimension (inhaltlich) mit der Bezeichnung ‚Vorläufigkeit‘ die Theoriedimensionen Alternativität, Transparenz, Metakommunikation, Begründungspflicht und Dynamik. Prominent sind hier Items und Formulierungen, die auf die „Vorläufigkeit allen Erkennens“ (Cramer et al., 2019, S. 415) hinweisen, wie sie idealtypisch im theoretischen Prinzip der Alternativität repräsentiert ist. Eine Aufmerksamkeit für diese Vorläufigkeit ist für Lehrpersonen relevant, um mit Ungewissheit adäquat umgehen zu können, die aus meta-reflexiver Perspektive als konstitutiv für das komplexe Handlungsfeld Schule angenommen wird. Aus der Vorläufigkeit resultiert eine Aufforderung an die (didaktische) Gestaltung des Lehramtsstudiums. Dozierende sollten diese Vorläufigkeit immer wieder thematisieren, ihr eigenes Vorgehen transparent machen und zugleich die Komplexität schrittweise erhöhen, um eine Überforderung zu Beginn des Studiums zu vermeiden. Konform zum Diskurs um *Personal Epistemology* manifestiert sich eine solche Position der Vorläufigkeit am ehesten in einer postrelativistischen Überzeugung, wonach die Gültigkeit von Erkenntnissen immer von der Prüfung theoretischer Argumente bzw. empirischer Evidenz abhängt und damit *per se* vorläufig ist. Im Gegensatz zu einer relativistischen Überzeugung ist damit nicht eine Beliebigkeit dieser Prüfung gemeint und Wissen legitimiert sich nicht allein durch die Übereinstimmung mit der eigenen Meinung. Daher überrascht die signifikante, positive Korrelation zwischen der Zustimmung zu postrelativistischen Aussagen und der Bedeutsamkeitseinschätzung meta-reflexiver Prinzipien dieser Dimension nicht, während eine Korrelation mit den beiden anderen Entwicklungsstufen epistemologischer Überzeugungen nicht signifikant ausfällt. Der Begriff ‚Vorläufigkeit‘ wird somit zur Beschreibung dieser Dimension als semantisch und konstruktbezogen angemessen erachtet.

Multiperspektivität: Nach Fives & Buehl (2012) besitzen Überzeugungen eine Filterfunktion für die Wahrnehmung, eine Rahmungsfunktion für die Bewertung und eine Steuerungsfunktion für das Handeln. Merk (2020) nennt epistemische Überzeugungen als Beispiel für die Rahmungsfunktion. Eine Lehrperson mit stark ausgeprägten absolutistischen Überzeugungen beurteilt demnach eine konkrete Situation anders als eine zweite Lehrperson mit stark ausgeprägten postrelativistischen Überzeugungen. Die zweite Lehrperson dürfte mit einer höheren Wahrscheinlichkeit adäquate Deutungen des komplexen

Handlungsfeldes vornehmen können. Die Analyse der Pilotierungen 3 und 4 weist den zweiten Faktor als denjenigen aus, der genau diesen Zusammenhang zwischen Meta-Reflexivität und Epistemischen Überzeugungen zu beschreiben scheint. Dieser Faktor kann als Multiperspektivität bezeichnet werden. Eine Person mit niedriger Merkmalsausprägung in dieser meta-reflexiven Dimension muss in der Professionalisierungsphase unterschiedliche Perspektiven auf einen Gegenstand kennen lernen (Dimension Mehrperspektivität) und verstehen, dass diese unterschiedlichen Perspektiven unabhängig voneinander ihre je spezifische Berechtigung haben können, was auf das meta-reflexive Prinzip der Unabhängigkeit verweist. Ebenso scheint das Prinzip Distanzierung auf, das auf die Bedeutsamkeit einer kritischen Grundhaltung, Distanz und Prüfung der vorgestellten Inhalte (vgl. Tab. 3) hinweist.

Erkenntnisgenese: Der dritte Faktor vereint die meta-reflexiven Prinzipien Kontextgebundenheit und Historisierung. Hat eine Person auf den bislang vorgestellten Faktoren hohe Merkmalsausprägungen, misst sie dem Einnehmen unterschiedlicher Perspektiven Bedeutung zu, ohne diese Unterschiede gleich nivellierend integrieren zu wollen. Dieser Faktor ergänzt die Bedeutsamkeitszuschreibung einer Person an die jeweiligen axiomatischen Grundlagen sowie historischen Kontexte von Theorie und Empirie. Eine entsprechende Verortung schließt aus Perspektive der Meta-Reflexivität auch die Kenntnis unterschiedlicher wissenschaftstheoretischer Paradigmen ein (vgl. Cramer et al., 2023, S. 10) sowie die Fähigkeit, sich mit diesen auf abstraktem Niveau auseinandersetzen zu können. Eine absolutistische Überzeugung würde sich in einer a priori Bevorzugung eines bestimmten, als wahr angesehenen Paradigmas niederschlagen. Dadurch wäre der Weg zu einer offenen Auseinandersetzung mit der Genese von Wissen mit Blick auf dessen axiomatische und historische Grundlagen verstellt. Eine postrelativistische Position wäre gegeben, wenn eine Aussage nur im Rahmen des jeweiligen axiomatischen bzw. zeithistorischen Kontextes als gültig angesehen wird. Somit erscheint die Bezeichnung des dritten Faktors mit ‚Erkenntnisgenese‘ angemessen. Dies legt auch die Analyse der Richtung der korrelativen Zusammenhänge zwischen diesem Faktor und den Ausprägungen epistemologischer Überzeugungen nahe (vgl. Tab. 3).

Die in Pilotierung 4 hervorgebrachte Skala zur Erfassung *der zugeschriebenen Bedeutsamkeit an meta-reflexive Prinzipien* für die Professionalisierung von (angehenden) Lehrpersonen ist das Ergebnis einer explorativen Faktorenanalyse, was als methodische Limitation anzusehen ist. In einer weiterführenden Untersuchung muss sich die Skala demnach noch konfirmatorisch bewähren. Da eine PFA anstelle einer Maximum-Likelihood-Faktorenanalyse durchgeführt wurde, konnte das gefundene Modell zudem nicht inferenzstatistisch überprüft werden, auch wenn die Matrix der Residuen zwischen empirischen und modellimplizierten Korrelationen auf einen guten Modellfit hindeutet. Das Durchführen der PFA erscheint jedoch adäquater für die vorliegenden Daten,

da sich keine multivariate Normalverteilung bezüglich des Antwortverhaltens zeigte. Die Skala basiert auf Erhebungen in Tübingen und Freiburg mit Lehramtsstudierenden in frühen Fachsemestern und wurde in Lehrveranstaltungen durchgeführt, die explizit meta-reflexiv angelegt sind. Ob die meta-reflexive Ausrichtung der Lehrveranstaltungen als Vor- oder Nachteil anzusehen ist, muss diskutiert werden: Außer Frage steht, dass die Skala auch an anderen Standorten im In- und Ausland sowie an verschiedenen Altersgruppen eingesetzt werden muss, um als funktional auch für diese Populationen gelten zu können. Die Teilnehmenden der Pilotierungen stehen am Anfang ihres Studiums und damit auch am Beginn ihrer eigenen Professionalisierung und dürften damit vielfach erstmals mit wissenschaftlichen Inhalten konfrontiert werden. Das Einbeziehen von älteren Studierenden, Referendar:innen sowie aktiven Lehrer:innen könnte Erkenntnisse darüber erlauben, wie sich die meta-reflexiven Bedeutsamkeitszuschreibungen im Verlauf der Professionalisierung entwickeln. Dadurch ergäbe sich auch eine größere Stichprobe, was die Angemessenheit der inferenzstatistischen Schlüsse erhöhen würde. Bislang ist mit dem FREE-EDPSY nur ein naheliegendes Instrument zur konvergenten Validierung der Skala mit eingesetzt worden. Das Hinzufügen weiterer Kovariaten könnte die Validität der Skala noch erhöhen, insbesondere ein Instrument als Maß für diskriminante Validität wäre in dieser Hinsicht nützlich. Die hier vorgestellte Operationalisierung nimmt nur einen Teilbereich von Meta-Reflexivität empirisch in den Blick und versucht dabei, so theoriekonform wie möglich zu sein. Herausforderungen beim Austarieren einer theoriekonformen Skalenentwicklung und deren empirischen Validierung zeigen sich bei jeder Pilotierung insbesondere im Prozess der Itemauswahl.

Unabhängig von der Skalenvalidierung sollte überprüft werden, ob die theoretisch postulierten meta-reflexiven Prinzipien bzw. die extrahierten Faktoren kognitiv repräsentiert sind und nicht beispielsweise auf Effekten sozialer Erwünschtheit beruhen. Dies zu eruieren, erfordert eine qualitativ-empirische Herangehensweise, in der beispielsweise Vignetten zu den gefundenen Faktoren entwickelt werden, zu welchen sich Proband:innen in kognitiven Interviews verhalten. Sollten sich die vorgestellten inhaltlichen Interpretationen der Faktoren durch die oben angeführten, ausstehenden skalenanalytischen Validierungsschritte bestätigen lassen und ein Hinweis auf deren kognitive Repräsentanz möglich werden, dann könnte auch über die Frage befunden werden, ob dieses spezifische meta-reflexive Konstrukt einer bestimmten Konstruktklasse – etwa Überzeugungen, Werthaltungen, Zielen oder Einstellungen – zugeordnet werden kann (vgl. Merk & Schmidt, 2023).

Im Ergebnis liegt infolge der Pilotierungen ein empirisch tragfähiges, validiertes und forschungsökonomisch handhabbares Instrument zur Erfassung der Bedeutsamkeitszuschreibungen an meta-reflexive Prinzipien in der Lehrer:innenbildung vor, mit dem weitere empirische Forschung zur Meta-Reflexi-

vität möglich wird. In Folgestudien kann die Faktorenstruktur konfirmiert und dann das Instrument für inhaltliche Fragestellungen eingesetzt werden, etwa indem diese Einschätzungen zur Bedeutsamkeit von Meta-Reflexivität in Beziehung zu anderen Variablen in der empirischen Lehrer:innenbildungsforschung gesetzt werden (zu personenbezogener Forschung vgl. etwa die Ausführungen in: Cramer et al., 2020). So wird ein wichtiger meta-reflexiver Schritt gegangen, nämlich die Bearbeitung eines Erkenntnisinteresses bzw. einer Fragestellung aus mehreren Perspektiven heraus (hier aus einer empirischen im unmittelbaren Anschluss an eine theoretische). Damit wird die professionstheoretische Perspektive der Meta-Reflexivität zunehmend zu einem *professionstheoretischen Ansatz* ausgearbeitet, indem die postulierte Relevanz von Meta-Reflexivität für die Lehrer:innenbildung und das Lehrpersonenhandeln einer empirischen Untersuchung zugänglich wird.

Anmerkung

Wir danken Sarah Forster-Heinzer, Jana Groß Ophoff, Martin Harant, Samuel Merk, Tom Rosman, Felix Schreiber und Elmar Teufl für wertvolle Hinweise zu unserem Vorgehen.

Literatur

- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Springer. <https://doi.org/10.1007/978-3-540-33306-7>
- Brandt, H. (2020). Exploratorische Faktorenanalyse (EFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 575–614). Springer. https://doi.org/10.1007/978-3-662-61532-4_23
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Cramer, C. (2020). Meta-Reflexivität in der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 204–214). Klinkhardt, utb. <https://doi.org/10.35468/hblb2020-024>
- Cramer, C. (2020). Professionstheorien. Überblick, Entwicklung und Kritik. In M. Harant, U. Kuchler & P. Thomas (Hrsg.), *Theorien!* (S. 111–128). Tübingen University Press. <http://dx.doi.org/10.15496/publikation-45602>
- Cramer, C., Groß Ophoff, J., & Schreiber, F. (2023). Professionalität im Lehrerinnen- und Lehrerberuf. Zur Inkohärenz von Professionstheorie und Überzeugungen von Lehrpersonen. *Zeitschrift für Pädagogik*, 69(1), 67–87. <https://doi.org/10.3262/ZP2301067>

- Cramer, C., Harant, M., Merk, S., Drahmman, M., & Emmerich, M. (2019). Meta-Reflexivität und Professionalität im Lehrerinnen- und Lehrerberuf. *Zeitschrift für Pädagogik*, 65(3), 401–423. <https://doi.org/10.3262/ZP1903401>
- Cramer, C., Brown, C., & Aldridge, D. (2023). Meta-Reflexivity and Teacher Professionalism: Facilitating Multiparadigmatic Teacher Education to Achieve a Future-Proof Profession. *Journal of Teacher Education*, 74(5), 467–480. <https://doi.org/10.1177/00224871231162295>
- Cramer, C., Rothland, M., König, J., & Blömeke, S. (Hrsg.) (2020). *Handbuch Lehrerinnen- und Lehrerbildung*. Klinkhardt, utb. <https://doi.org/10.35468/hblb2020>
- Cudek, R. (2000). Exploratory factor analysis. In H. Tinsley & S. D. Brown (Hrsg.), *Handbook of Applied multivariate statistics and mathematical modelling* (S. 265–296). Academic press.
- Dinno, A. (2009). Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*, 44(3), 362–388. <https://doi.org/10.1080/00273170902938969>
- Eid, M., Gollwitzer, M., & Schmitt, M. (2017). *Statistik und Forschungsmethoden* (5. Aufl.). Beltz.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the „messy“ construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer & M. Zeidner (Hrsg.), *APA educational psychology handbook. Vol. 2.* (S. 471–499). American Psychological Association. <https://doi.org/10.1037/13274-019>
- Gäde, J. C., Schermelleh-Engel, K., & Brandt, H. (2020). Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 615–659). Springer. https://doi.org/10.1007/978-3-662-61532-4_24
- Hartmann, U., Kindlinger, M., & Trempler, K. (2021). Integrating information from multiple texts relates to pre-service teachers' epistemic products for reflective teaching practice. *Teaching and Teacher Education*, 97(2), 209–220. <https://doi.org/10.1037/a0023353>
- Helsper, W. (2021). *Professionalität und Professionalisierung in pädagogischen Handlungsfeldern: Eine Einführung*. Budrich, utb.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185. <https://doi.org/10.1007/BF02289447>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling*. R package version 0.5-6. <https://CRAN.R-project.org/package=semTools>
- Kindlinger, M. (2021). Meta-reflexivity and epistemic cognition in social science teacher education. *Journal of Social Science Education*, 20(3), 29–54. <https://doi.org/10.11576/jsse-4087>
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). „MVN: An R Package for Assessing Multivariate Normality.“ *The R Journal*, 6(2), 151–162. <https://journal.r-project.org/archive/2014-2/korkmaz-goksuluk-zararsiz.pdf>

- Krettenauer, T. (2005). Die Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen und das Problem der Übertragbarkeit von Interviewverfahren in standardisierte Fragebogenmethoden. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 37(2), 69–79. <https://doi.org/10.1026/0049-8637.37.2.69>
- Merk, S. (2020). Überzeugungen. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 825–832). Klinkhardt, utb. <https://doi.org/10.35468/hblb2020-102>
- Merk, S., & Schmidt, K. (2023). Grundlegende Fragen an eine quantitativ-empirische Erfassung von Meta-Reflexivität. In C. Cramer (Hrsg.), *Meta-Reflexivität und Professionalität von Lehrpersonen. Theorieentwicklung und Forschungsperspektiven* (S. 143–154). Waxmann.
- Merk, S., Schneider, J., Bohl, T., Kelava, A., & Syring, M. (2017). Epistemologische Überzeugungen von Lehramtsstudierenden bezüglich pädagogischen Wissens: Gegenstands-, Quellen- und Kontextspezifität. *Journal for Educational Research Online*, 9(1), 169–189. <https://doi.org/10.25656/01:12973>
- Moosbrugger, H., & Brandt, H. (2020). Itemkonstruktion und Antwortverhalten. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 67–89). Springer. https://doi.org/10.1007/978-3-662-61532-4_4
- Revelle, W. (2023). *psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.3.3*. <https://CRAN.R-project.org/package=psych>
- Robitzsch, A. (2020). Why Ordinal Variables Can (Almost) Always Be Treated as Continuous Variables: Clarifying Assumptions of Robust Continuous and Ordinal Factor Analysis Estimation Methods. *Frontiers in Education*, 5, 177. <https://doi.org/10.3389/educ.2020.589965>
- Rosman, T., Mayer, A.-K., Merk, S., & Kerwer, M. (2019). On the benefits of ‚doing science‘: Does integrative writing about scientific controversies foster epistemic beliefs? *Contemporary Educational Psychology*, 58, 85–101. <https://doi.org/10.1016/j.cedpsych.2019.02.007>
- Rosman, T., Schlag, M., & Merk, S. (2020). Das Zusammenspiel epistemischer Überzeugungen und der Bedeutsamkeitseinschätzung pädagogisch-psychologischen Wissens im Lehramtsstudium: Längsschnittliche und querschnittliche Analysen. *Psychologie in Erziehung und Unterricht*, 67(3), 164–177. <https://doi.org/10.2378/peu2020.art15d>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://www.jstatsoft.org/v48/i02/>
- Rossiter, J. R. (2008). Content Validity of Measures of Abstract Constructs in Management and Organizational Research. *British Journal of Management*, 19(4), 380–388. <https://doi.org/10.1111/j.1467-8551.2008.00587.x>
- Schermelleh-Engel, K., & Gäde, J. C. (2020). Modellbasierte Methoden der Reliabilitätsschätzung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 335–368). Springer. https://doi.org/10.1007/978-3-662-61532-4_15
- Skott, J. (2015). The promises, problems, and prospects of research on teachers' beliefs. In H. Fives & M. G. Gill (Hrsg.), *International handbook of research on teachers' beliefs. Vol. 1* (S. 13–30). Routledge.

- Thiel, C. (2018). Art. Theorie. In J. r. Mittelstraß (Hrsg.), *Enzyklopädie Philosophie und Wissenschaftstheorie. Band 8 (Th–Z)* (S. 20–29). Metzler.
- Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*(2), 209–220. <https://doi.org/10.1037/a0023353>
- Wilde, A., & Kunter, M. (2016). Überzeugungen von Lehrerinnen und Lehrern. In M. Rothland (Hrsg.), *Beruf Lehrer/Lehrerin. Ein Studienbuch* (S. 299–315). Waxmann.
- Yousfi, S. (2005). Mythen und Paradoxien der klassischen Testtheorie (II). Trennschärfe und Gütekriterien. *Diagnostica, 51*(2), 55–66. <https://doi.org/10.1026/0012-1924.51.2.55>