

# Grundlagen der Kompetenzfeststellung

## Handlungs- und kompetenzorientiert Leistungen messen und bewerten



### Kurz zusammengefasst ...

Der Artikel beleuchtet die Rolle von Prüfungen und Kompetenzmessungen in der handlungsorientierten Feuerwehrausbildung. Er zeigt Funktionen, Grenzen und Gütekriterien von Prüfungen auf und stellt Status- und Prozessdiagnostik gegenüber. Anhand praxisnaher Beispiele werden verschiedene Prüfungsformate diskutiert und deren Beitrag zur validen Erfassung fachlicher und überfachlicher Kompetenzen aufgezeigt. Der Beitrag bietet konkrete Orientierung für eine zeitgemäße, lernförderliche Prüfungsgestaltung.



In einer handlungsorientierten Feuerwehrausbildung sollen Teilnehmerinnen und Teilnehmer Kompetenzen in verschiedenen Bereichen und Niveaus erweitern.

Die Auswahl der Kompetenzen erfolgt anhand der Vorgaben des jeweiligen Qualifikationsprofils und berücksichtigt ebenfalls die individuellen Eingangskompetenzen der Teilnehmenden. Somit wird sichergestellt, dass Teilnehmende genau die Kompetenzen erlernen, die für sie und ihre künftige Aufgabe relevant sind. Zum Abgleich während des Lernens und zur Bestätigung des Lernerfolges werden während der Ausbildung Kompetenzmessungen durchgeführt.

### Bedeutung, Funktionen und Grenzen von Prüfungen

Prüfungen sind als Werkzeug der Normierung ein bedeutendes Instrument im Lernprozess. Sie geben Auskunft über Lernerfolge und Leistungsdefizite. In der Feuerwehrausbildung haben Prüfungen eine hervorstechende Bedeutung erlangt, da sie das entscheidende Kriterium über den Erfolg

oder Misserfolg einer Ausveranstaltung symbolisieren. Durch diese Funktion beeinflussen Prüfungen den Lernweg der Teilnehmenden enorm. Es wird gelernt, was prüfungsrelevant ist. Ausbilderinnen und Ausbilder potenzieren diese Verzerrung, indem sie prüfungsrelevante Inhalte besonders hervorheben und ihnen überproportional Zeit einräumen. Beachten wir, dass aktuelle Prüfungsformen häufig eine geringe Aussagekraft (Gütekriterium: Validität) für die eigentlich intendierte Aufgabe besitzen, so wird klar, dass Prüfungen in der Feuerwehrausbildung dringend an die aktuellen Anforderungen angepasst werden sollten. Außerdem gilt zu beachten, dass wenn Lernprozesse verändert werden, Prüfungsformen und -inhalte in diesem Prozess ebenso angepasst werden müssen. Ein veraltetes Prüfungsdesign bremst die Umstellung auf moderne Ausbildungsvarianten und erzeugt Irritationen und Diskussionen bei Teilnehmenden und Ausbildenden.

Bei einer Kompetenzmessung wird in einem festgeschriebenen Setting ein beobachtetes Verhalten als Ist-Wert mit einem Soll-Wert verglichen und dessen Abweichung vermerkt. Dabei ist zu beachten, dass es sich



dabei stets um eine Stichprobe handelt, die zudem nur einen Bruchteil der real vorhandenen Kompetenzen darstellt. Begutachtet werden kann lediglich das gezeigte Verhalten, also die Performanz, in der jeweiligen Prüfungssituation. Dadurch bedingt ist die Aussagekraft von Prüfungen limitiert. Dennoch besitzen Prüfungen Funktionen, die den Lernprozess als Ganzes unterstützen. Neben der Normierung des individuellen Lernerfolgs ist hier insbesondere die Rückmeldefunktion hervorzuheben. Unter der Normierung wird verstanden, dass der individuelle Lernerfolg der Teilnehmenden einem Wert zugewiesen wird. Damit kann ein Modul oder Lehrgang als abgeschlossen gekennzeichnet werden. Dies wird insbesondere genutzt, um Qualifikationen auszusprechen (z.B. Qualifikation Atemschutzgeräteträger).

Die Funktion der Rückmeldung ist insbesondere für die Lernenden bedeutsam. Ein konkretes Feedback hilft Teilnehmenden, ihre gezeigten Leistungen einzuordnen, Erfolge zu kennzeichnen und Defizite klar zu benennen. Das stärkt die Selbstwirksamkeit und steigert die intrinsische Motivation. Ein persönliches Feedback gehört zu den stärksten Faktoren, eine Lernkurve positiv zu beeinflussen (Hattie & Timperley 2007). Prüfungsergebnisse sollen daher ungefiltert an die Teilnehmenden weitergegeben werden.

Neben den Lernenden können auch Ausbildende und Ausbildungsorganisierende gezeigte Leistungen als Indikator für den Erfolg einer Ausbildungsveranstaltung nutzen. Durch eine systematische Auswertung können kontinuierliche Verbesserungen auf validen Ergebnissen basieren.

Wie bereits beschrieben, gilt es zu beachten, dass eine Kompetenzmessung, egal wie umfangreich sie ausfällt, lediglich eine Stichprobe ist. Gemessen wird nur zu einem

bestimmten Zeitpunkt in einer konkreten Situation. Mit dem Verständnis, dass Kompetenzen als Bereitschaft des Handelns verstanden werden, kann in einer Prüfung lediglich die Performanz in Bezug auf die Aufgabenstellung sichtbar werden. Die Messung liefert einen Indikator für das Lernen, ein Rückschluss auf die Gesamtheit des Lernprozesses ist nur sehr bedingt möglich.

**Prüfungen sind ein sehr bedeutender Teil des Lernprozesses und müssen daher überlegt in diesen eingebaut sein.**

In der handlungsorientierten Feuerwehrausbildung wird die Entwicklung umfassender Handlungskompetenzen als übergeordnetes Ausbildungsziel festgeschrieben. Handlungskompetenz wird verstanden als die Bereitschaft und Befähigung, in wechselnden Situationen sachgerecht durchdacht sowie individuell und sozial verantwortlich zu handeln. Neben fachlichen Anforderungen stehen damit auch soziale und individuelle Fähigkeiten im Anforderungsprofil. Um dieses Ausbildungsziel sichtbar abbilden zu können, müssen innerhalb der Kompetenzmessung neben der Fachkompetenz auch Teamwork, Kommunikation und Problemlösefähigkeit als Stellvertreter der überfachlichen Kompetenzen berücksichtigt werden. Damit gilt, dass für die Feuerwehrausbildung Prüfungsdesigns anzuwenden sind, in denen fachliche und überfachliche Kompetenzen sichtbar werden und eine gemeinsame Begutachtung möglich wird. Übliche Prüfungsdesigns werden zum Abschluss dieses Artikels beschrieben.

### Gütekriterien

Für Prüfungen sind die drei Gütekriterien Objektivität, Validität und Reliabilität maßgeblich.

Die Objektivität eines Verfahrens bezeichnet den Grad, in dem ein

Untersuchungsergebnis unabhängig vom Untersucher ist (Hesse & Latzko, 2011). Ein vollkommen objektives Mess-verfahren ist demnach gegeben, wenn verschiedene Untersuchende zu exakt gleichen Ergebnissen gelangen, beispielsweise wenn mehrere Prüfende in einem Führungslehrgang bei einer bewerteten Einsatzübung anhand von vorher festgelegten Kriterien dasselbe Ergebnis feststellen. Die Objektivität kann in die drei Phasen Durchführungsobjektivität, Auswertungs-objektivität und Interpretationsobjektivität gegliedert werden.

Die Validität eines Verfahrens bezeichnet einen Grad der Zuverlässigkeit, mit dem ein Verfahren das Merkmal, das es messen soll oder zu messen vorgibt, auch tatsächlich misst (Hesse & Latzko, 2011). Ein Messverfahren ist demnach vollkommen valide, wenn das Ergebnis einen unmittelbaren Rückschluss auf den Ausprägungsgrad des zu erfassenden Merkmals zulässt. Soll in einem Gerätewart-Lehrgang beurteilt werden, inwieweit die Teilnehmenden in der Lage sind, eine Leiter auf ihre Funktionsfähigkeit zu überprüfen, lässt sich die Aussagekraft erhöhen, indem sie diese Funktionsprüfung praktisch zeigen anstatt auswendiggelernte Prüf-daten wiederzugeben. Die Validität ist auch dann geringer, wenn Einflüsse auf das Messverfahren einwirken, die eigentlich nicht erfasst werden sollen. So kann es die Aussagekraft einer Prüfung beispielsweise auch mindern, wenn Teilnehmende unter Prüfungsangst oder Zeitdruck vor den Ausbildenden stehen (Frenzel, Götz, Pekrun, 2015).

Die Reliabilität eines Verfahrens bezeichnet den Grad der Genauigkeit, mit dem ein Verfahren ein Merkmal misst – unabhängig davon, was es zu messen vorgibt (Hesse & Latzko, 2011). Die Reliabilität eines Messverfahrens kann demnach bestimmt werden, indem bei wiederholenden Messungen

derselben Messobjekte und anschließendem Vergleich der einzelnen Messergebnisse diese übereinstimmen.

Grundsätzlich gilt es, diese Gütekriterien gemeinsam zu berücksichtigen. Innerhalb der Feuerwehrausbildung kann jedoch konstatiert werden, dass insbesondere das Gütekriterium der Validität in der Vergangenheit unzureichend berücksichtigt wurde. Der Versuch, möglichst objektive und reliable Prüfungen zu erstellen, ergab Multiple-Choice-Prüfungen, in denen deklaratives Wissen zugeordnet wurde. Die Validität ging dabei gänzlich verloren. Wird beispielsweise in einer Multiple-Choice-Prüfung der wortgenaue Inhalt von Gesetzestexten abgefragt, lässt sich daraus nicht ableiten, ob die Teilnehmenden tatsächlich in der Lage sind, nach diesen Vorschriften zu handeln. Es ist offensichtlich, dass die Aussagekraft von gezeigten Leistungen in unserer Tätigkeit das Kriterium ist, an dem wir uns im Einsatzgeschehen gerne messen lassen. Demnach sollte das Gütekriterium der Validität zukünftig in den Fokus rücken und eine besondere Bedeutung erlangen.

**Zur Steigerung der Aussagekraft einer Prüfung sollten Auszubildende kritisch prüfen, ob relevante Kompetenzen erfasst werden.**

### **Kompetenzmessung im Lernprozess**

Während der Ausbildung ist es Aufgabe der Ausbildenden, den Lernprozess zu steuern und diesen an die Vorgaben und den Leistungsstand der Lernenden anzupassen. Daher ist es bedeutsam, dass Kompetenzmessungen nicht nur einmalig am Ende eines Ausbildungsabschnittes erfolgen, sondern bestenfalls kontinuierlich. Dabei kann den Lernenden rückgemeldet werden, was von ihnen erreicht wurde und welche Verhaltensweisen dabei für den Lern-erfolg maßgeblich waren. Der Beobachtungsfokus kann bei der Lernprozess-Messung auf



Handlungsabläufe, Strategien, Lösungswege und Reflexionsansätze gelegt werden. Damit besitzt die Prozessdiagnostik - eine den Lernprozess begleitende Messung - einen bedeuten-den Mehrwert zu einer Statusdiagnostik, die lediglich den Erfolg einer Ausbildung am Ende bestimmt.

Regelmäßige Rückmeldungen geben den Teilnehmenden Orientierung und helfen Herausforderungen anzunehmen. Es ist möglich, dabei Werkzeuge wie die Selbstreflexion sowie das Peergroup-Feedback einzubeziehen. Rückmeldungen aus der Reihe der Teilnehmenden besitzen häufig einen besonderen Wert für die Lernenden und können daher einen bedeutsamen Mehrwert liefern. Eine Selbsteinschätzung hilft oft, die Verantwortung für den eigenen Lernprozess zu verinnerlichen.

Ein Feedback von Ausbildenden soll dabei nicht automatisch und allein die Funktion einer Bewertung mit Einfluss auf die abschließende Beurteilung haben. Insbesondere im Sinne eines zeitgemäßen Umgangs mit Fehlern und einer Methodik, die Ausprobieren und eigenständiges Entdecken beinhaltet, sollen Fehler als Lernchancen gesehen werden. Reine Leistungssituationen sollten daher stets als solche kenntlich gemacht werden. Eine Vermischung von Lern- und Leistungssituationen kann mit einem ungünstigen, negativen Lernklima einhergehen (Hascher, Götz, Zirngibl & Pekrun, 2004). Absolviert beispielsweise eine Teilnehmerin im Verlauf eines Zugführerlehrgangs eine Planübung als Training, kann sie gehemmt sein, Fehler zu machen, wenn sie befürchtet, dass diese sich negativ auf ihre abschließende Bewertung auswirken.

Empfehlenswert ist, dass in den Phasen Information und Planung die Teilnehmenden bewusst entdeckend, testend und selbstgesteuert unterwegs sind. Mit Abschluss der

Phase Durchführung kann eine eindeutig gekennzeichnete Leistungssituation mit transparentem Anforderungsprofil die Grundlage eines abschließenden Feedbacks liefern.

Durch einen regelmäßigen Abgleich der tatsächlich erbrachten Leistungen mit den in der Planung fixierten Zielwerten kann der Lernprozess zielgerichtet gesteuert werden.

Statusdiagnostik bezeichnet eine Form der Leistungsmessung, bei der die Leistung der Prüfenden zu einem bestimmten Zeitpunkt erfasst wird. Ziel ist es, qualifizierende Bewertungen vorzunehmen oder festzustellen, ob ein definiertes Kompetenzniveau erreicht wurde (Looney 2011). Der Fokus liegt dabei auf dem Endstand des Lernprozesses und nicht auf dessen Verlauf.

Im Kontext der Feuerwehrausbildung dient Statusdiagnostik vor allem dazu, die Berechtigung zu erlangen, bestimmte Ämter zu übernehmen oder festgelegte Funktionen im Einsatz auszuführen. Die Leistungsmessung ist somit eng mit formalen Qualifikationen und Einsatzbefugnissen verknüpft.

Charakteristisch für diese Form der Leistungsmessung ist der starke Fokus auf ein konkretes Leistungsprodukt. Beispiele hierfür sind ausgefüllte Prüfprotokolle, die korrekt durchgeführte Erkundung und Befehlsgebung durch eine Führungskraft oder die Bearbeitung einer standardisierten Textaufgabe. Klassischerweise sind die Prüfungsszenarien am Ende eines Lehrgangs sehr genau vorgegeben. Durch klar definierte Kriterien soll eine möglichst eindeutige Bewertung und eine hohe Vergleichbarkeit zwischen den Prüflingen erreicht werden.

Diese starke Standardisierung birgt jedoch auch Risiken. Wird ausschließlich ein punktuelles Leistungsprodukt betrachtet, lassen sich zentrale Kompetenzen wie

Problemlösefähigkeit, Planungsfähigkeit, Selbstkorrektur im laufenden Einsatzgeschehen, der gewählte Lösungsweg oder die Begründung von Entscheidungen nur unzureichend erfassen. Insbesondere dann, wenn isoliertes, von der praktischen Anwendung losgelöstes Wissen abgeprüft wird, besteht die Gefahr einer geringen Validität der Leistungsmessung.

Um dem entgegenzuwirken, sind möglichst realitätsnahe Prüfungsszenarien erforderlich, in denen die erwarteten Kompetenzen aussagekräftig gemessen werden können.

Darüber hinaus finden überfachliche Kompetenzen wie Eigenverantwortung, Teamarbeit und zielgerichtetes Arbeiten in abschließenden, qualifizierenden Leistungsmessungen der Feuerwehr bislang nur wenig oder keine systematische Berücksichtigung. Um das überfachliche Leistungen ebenfalls beurteilen zu können, sind dafür strukturierte, Beobachtungs- und Bewertungsbögen notwendig, die eine nachvollziehbare und möglichst objektive Einschätzung ermöglichen.

Eine abschließende Beurteilung zur Qualifizierung von Einsatzkräften sollte in möglichst realitätsnahen, beobachtbaren Settings und anhand von aussagekräftigen Kriterien durchgeführt werden.

**Eine Beurteilung zur Qualifizierung von Einsatzkräften sollte in realitätsnahen Settings anhand von aussagekräftigen Kriterien durchgeführt werden.**

### **Feedbacksystem**

Ein Feedbacksystem ist ein methodisches Gerüst, in dem Feedback-Gespräche in einem strukturierten Rahmen wiederkehrend und aufeinander aufbauend erfolgen.

Grundsätzlich gilt, dass jede Art der Rückmeldung in einem für die Lernenden anknüpfbarem Zeitfenster erfolgt und in einem geschützten Setting stattfindet. Da das Feedback ein Werkzeug ist, welches das Verhalten in einer konkreten Situation in einem anschließenden Gespräch aufgreift, finden diese Gespräche fast ausnahmslos mit nur einem Lernendem statt. Diese Individualität gepaart mit der konkreten Situation ermöglicht einen Leistungsabgleich, der dem Lernenden sehr präzise gespiegelt werden kann.

Damit das Feedback den Gütekriterien entspricht und seine volle Wirksamkeit entfalten kann, ist es wichtig, dass im Gespräch Beobachtungen klar von Bewertungen getrennt werden. Die Wiedergabe von beobachtetem Verhalten, welches vom Lernenden als veränderbar eingeschätzt wird, ist der Grundstein hierfür. Hierbei sollen selbstverständlich sowohl Leistungserfolge, wie auch Leistungsdefizite angesprochen werden. Anschließend können für Leistungsdefizite Veränderungsoptionen aufgezeigt oder gemeinsam entwickelt werden. Zum Teil kann es helfen, wenn die Leistung des Lernenden im Vergleich zu vorher definierten Qualitätsstufen, die maßgeblich zur Erreichung des Erfolges der Ausbildungsveranstaltung sind, transparent dargestellt wird.

Einem Feedback vorangehen sollte eine Selbsteinschätzung, in dem die Lernenden ihr eigenes Planen, Entscheiden und Handeln reflektieren und so Verantwortung für die eigene Lernentwicklung übernehmen.

### **Peergroup-Feedback**

Beim Peergroup-Feedback geben Teilnehmende einem Mitglied der Gruppe eine Rückmeldung, die auf vorher festgelegten Kriterien beruht. Ziel ist dabei, das Vorgehen der feedback-empfangenden Person



konstruktiv zu optimieren (vgl. Mattes 2011). Auch die anderen Mit-glieder der Gruppe können so durch Beobachtung lernen.

Für eine lernförderliche Umsetzung sind klare Feedbackregeln erforderlich. In einer ersten Runde sollten ausschließlich positive Aspekte benannt werden, um Stärken sichtbar zu machen und eine wertschätzende Atmosphäre zu schaffen. Auszubildende tragen die Verantwortung dafür, dass diese Regeln eingehalten werden und in der ersten Runde keine negativ konnotierte Kritik in positiv verpackter Form geäußert wird.

Kritische Rückmeldungen sollten stets als konkrete Verbesserungsvorschläge formuliert sein und sich auf klar benennbare, beobachtbare Aspekte beziehen. Zudem sollte die feedback-empfangende Person zunächst Gelegenheit zur Selbstreflexion erhalten.

Ist eine positive Feedback- und Besprechungskultur etabliert, können Peergroup-Feedbackrunden für alle Beteiligten zu konstruktiven und motivierenden Lernerfahrungen werden (vgl. Mattes 2011). Die Rückmeldung erfolgt auf Augenhöhe, da alle Beteiligten der-selben sozialen Rolle als Teilnehmende angehören. Diese Gleichrangigkeit kann dazu beitragen, dass Feedback als unterstützend und weniger bewertend wahrgenommen wird. In diesem Rahmen eignet sich das Format besonders für die Prozessdiagnostik, da Lernfortschritte und Entwicklungsbedarfe kontinuierlich sichtbar gemacht werden.

Peergroup-Feedback setzt jedoch eine tragfähige Vertrauensbasis voraus, ohne diese wird Feedback nicht angenommen oder bleibt oberflächlich. Unspezifische Rückmeldungen wie pauschales Lob sind für die Lernentwicklung nicht hilfreich. Wenn Auszubildende im Anschluss eigenes Feedback geben, besteht zudem die Gefahr, dass das

Peergroup-Feedback relativiert wird. Dem kann begegnet werden, indem sie ausschließlich neue Aspekte ergänzen. Für qualifizierende Statusdiagnostik ist Peergroup-Feedback ungeeignet, da Teilnehmende nicht abschließend über Ausbildungserfolg oder -misserfolg urteilen können.

### Praktische Prüfungen

Bereits jetzt nehmen praktische Prüfungen in der Feuerwehrausbildung eine zentrale Rolle ein, da sie es ermöglichen, Handlungskompetenz unter realitätsnahen Bedingungen festzustellen. Sie können in unterschiedlichen Formen durchgeführt werden, etwa als Stationsprüfungen für einen Trupp, als Einzelprüfungen in herausfordernden Problem- oder Entscheidungssituationen, als Einsatzübungen oder in Form von Planübungen und Planbesprechungen. Allen Formaten ist gemeinsam, dass sie das tatsächliche Handeln der Teilnehmenden in den Mittelpunkt stellen.

Ein wesentlicher Vorteil praktischer Prüfungen liegt darin, dass sowohl das Handeln einzelner Teilnehmender als auch das Zusammenwirken im Team beobachtbar wird. Entscheidungsfindung, Durchführung von Maßnahmen und die laufende Kontrolle des eigenen Handelns können unmittelbar nachvollzogen werden. Durch die hohe Realitätsnähe der Szenarien handeln die Teilnehmenden sehr nah an tatsächlichen Einsatzbedingungen und werden darin gefordert, Probleme eigenständig und situationsangemessen zu lösen. Dadurch lassen sich nicht nur fachliche Kompetenzen, sondern auch überfachliche Fähigkeiten wie Teamarbeit und Kommunikation erfassen. Dabei können auch nur Teile einer ganzen Einsatzübung durchgeführt werden, beispielsweise nur der Erkundungsvorgang mit der Befehlsgebung einer Führungskraft.

Insbesondere Einsatz- und Planübungen bieten die Möglichkeit, Prüfungen eng mit dem Lernprozess zu verknüpfen. Teilnehmende können Feedback sowohl von Ausbildenden als auch von anderen Teilnehmenden erhalten, was diese Formate besonders geeignet für eine prozessorientierte Diagnostik macht.

Neben diesen Vorteilen sind praktische Prüfungen jedoch auch mit Herausforderungen verbunden. Sie erfordern im Vergleich zu schriftlichen Prüfungen einen deutlich höheren Einsatz von Zeit und Personal, da die Teilnehmenden nicht gleichzeitig geprüft werden können. Zudem ist die Entwicklung eines klaren, anwendungspraktischen und kriteriengeleiteten Bewertungsbogens notwendig, um eine möglichst objektive Beurteilung sicherzustellen. Für die Prüfenden kann es eine besondere Herausforderung darstellen, fachliche und überfachliche Leistungen gleichzeitig zu beobachten und angemessen zu bewerten. Je realitätsnäher die Prüfungsszenarien gestaltet sind, desto wichtiger ist es außerdem, den Teilnehmenden im Vorfeld transparent zu vermitteln, welche Anforderungen gestellt werden und nach welchen Kriterien die Bewertung erfolgt.

### **Mündliche Prüfungen**

Mündliche Prüfungen eignen sich insbesondere für Prüfungsformate, in denen Ergebnispräsentationen, Reflexionsprozesse und kommunikative Kompetenzen im Vordergrund stehen. Darüber hinaus sind mündliche Prüfungen für Gesprächssituationen wie Rollenspiele geeignet, insbesondere dann, wenn Konfliktsituationen simuliert und kommunikative Fähigkeiten überprüft werden sollen.

In der Durchführung mündlicher Prüfungen können je nach Prüfungssituation ein oder zwei Prüflinge beteiligt sein; in jedem Fall

sollten mindestens zwei Prüfende anwesend sein, um ein Mindestmaß an Objektivität sicherzustellen. Zwingende Voraussetzung ist zudem ein im Vorfeld festgelegter Kriterien- oder Bewertungsbogen, der gewährleistet, dass genau die geforderten Kompetenzen geprüft werden. Ohne eine solche Struktur besteht die Gefahr, dass Prüfungsentscheidungen stark von subjektiven Einschätzungen geprägt sind.

Ein zentrales Potenzial mündlicher Prüfungen liegt darin, dass sie es den Prüfenden ermöglichen, durch gezielte Nachfragen und Interaktion ein tieferes Verständnis der Prüflinge sichtbar zu machen, anstatt lediglich reproduziertes Wissen zu erfassen. Mündliche Prüfungen bieten insbesondere dann einen Mehrwert, wenn sie Raum für Begründungen, Reflexionen und situatives Denken lassen (Nallaya et al. 2024), zum Beispiel als Einsatznachbesprechungen. In solchen Settings erhalten die Teilnehmenden die Gelegenheit, ihr eigenes Handeln kritisch zu analysieren, Entscheidungen zu begründen, Fehler zu erkennen und zu benennen, was sie in zukünftigen Situationen anders machen würden. Fehler werden auf diese Weise nicht als Defizite, sondern als wertvolle Lerngelegenheiten genutzt, aus denen nachhaltige Erkenntnisse entstehen können (Brünken, Münzer & Spinath 2019).

Die mündliche Prüfung kann hier als Werkzeug zur Selbstreflexion genutzt und mit Präsentationsformen kombiniert werden (zum Beispiel in Form einer Besprechung einer erstellten Lagekarte in der technischen Einsatzleitung).

Je dynamischer und offener eine Gesprächssituation gestaltet ist, desto schwieriger wird es, sicherzustellen, dass tatsächlich die vorgesehenen Kompetenzen geprüft werden. Umgekehrt kann eine sehr strikte Orientierung an vorgegebenen Fragen und



Kriterien die Möglichkeit einschränken, tiefergehendes Verständnis oder flexible Problemlösefähigkeit sichtbar zu machen. Zudem lässt sich Handlungskompetenz nur eingeschränkt in mündlichen Prüfungen erfassen, da sie letztlich die Beobachtung tatsächlichen Handelns erfordert.

Teilnehmende, die von ihrer Persönlichkeit her zurückhaltend sind oder negative Erfahrungen mit mündlichen Prüfungen gemacht haben, können ihre Kompetenzen in mündlichen Prüfungsformaten unter Umständen nicht angemessen zeigen. Wenn kommunikative Kompetenz nicht explizit Prüfungsgegenstand ist, können diese Faktoren die Validität der Prüfungsergebnisse verzerren.

Um dem entgegenzuwirken, sollten mündliche Prüfungen möglichst authentisch an tatsächlich relevante Kommunikationssituationen angelehnt werden. In der Feuerwehrpraxis können dies etwa Lagebesprechungen, strukturierte Einsatzreflexionen oder realitätsnahe Konfliktgespräche sein. Werden solche Formate konsequent an realen Anforderungen ausgerichtet, lässt sich die Aussagekraft mündlicher Prüfungen erhöhen und ihre Passung zu den angestrebten Kompetenzen verbessern.

### Schriftliche Prüfungen

Schriftliche Prüfungen sind geeignet, um anwendungsbezogenes, fachliches Wissen zu überprüfen. Darüber hinaus eignen sie sich auch dazu, ausgehend von konkreten Einsatzszenarien die Fähigkeiten der Teilnehmenden zum Planen, Entscheiden und Begründen und Reflektieren abzufragen. Schriftliche Prüfungen können auch digital durchgeführt werden und bieten damit eine hohe organisatorische Flexibilität.

Grundsätzlich gilt, dass Prüfungsformate mit offenen Fragen eine höhere Validität aufweisen und höhere kognitive

Verständnisniveaus erfassen können als geschlossene Formate (Mehta et al. 2016). Sie verlangen von den Teilnehmenden, eigene Antworten zu formulieren, Zusammenhänge darzustellen und Entscheidungen zu begründen, was Rückschlüsse auf ein tieferes Verständnis zulässt. Multiple-Choice-Prüfungen sind hingegen vor allem deshalb verbreitet, weil sie eine höhere Auswertungsobjektivität und Reliabilität bieten und sich schnell und einfach auswerten lassen. Werden Multiple-Choice-Aufgaben sorgfältig und anspruchsvoll konstruiert, können sie durchaus über reines Faktenwissen hinausgehen und auch komplexere Denkprozesse ansprechen (Mehta et al. 2016). Die Entwicklung solcher Aufgaben ist jedoch anspruchsvoll und erfordert einen hohen didaktischen und fachlichen Aufwand.

Eine besondere Chance schriftlicher Prüfungen liegt darin, dass die Teilnehmenden ihre Antworten in Ruhe planen, abwägen, begründen und reflektieren können, ohne dem unmittelbaren Zeit- oder Antwortdruck ausgesetzt zu sein, der in mündlichen Prüfungsformaten häufig entsteht. Dies kann insbesondere bei komplexen Fragestellungen zu differenzierten und reflektierten Antworten führen.

Wird in schriftlichen Prüfungen lediglich isoliertes Faktenwissen abgefragt, ist die Aussagekraft der Prüfung gering, da kaum Rückschlüsse auf die tatsächliche Handlungskompetenz möglich sind. Einsatzszenarien müssen daher sorgfältig und realitätsnah gestaltet werden, um sichtbar zu machen, wie Teilnehmende planen, entscheiden und ihre Maßnahmen begründen würden. Offene Aufgabenformate sind zudem vergleichsweise aufwendig in der Erstellung und bergen das Risiko einer geringeren Objektivität und Reliabilität in der Auswertung, sofern keine klaren,

kriteriengeleiteten Bewertungsmaßstäbe festgelegt werden.

**Die Auswahl der Prüfungsform erfolgt aufgrund der festzustellenden Kompetenzen.**

| Format                        | Prozessdiagnostik | Statusdiagnostik | Fachliche Kompetenzen sichtbar | Überfachliche Kompetenzen sichtbar |
|-------------------------------|-------------------|------------------|--------------------------------|------------------------------------|
| <b>Feedbacksystem</b>         | ja                | ja               | ja                             | ja                                 |
| <b>Peer-group-Feedback</b>    | Ja                | Ja               | Ja                             | ja                                 |
| <b>Selbsteinschätzung</b>     | Ja                | z.T.             | z.T.                           | ja                                 |
| <b>Praktische Prüfungen</b>   | z.T.              | ja               | ja                             | z.T.                               |
| <b>Mündliche Prüfungen</b>    | z.T.              | Ja               | z.T.                           | nein                               |
| <b>Schriftliche Prüfungen</b> | Nein              | Ja               | z.T.                           | nein                               |

## Literatur

Brünken, B., Münzer, S. & Spinath, B. (2019). Pädagogische Psychologie – Lernen und Lehren. Göttingen: Hogrefe.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. Review of Educational Research.

Hesse, I. & Latzko, B. (2011). Diagnostik für Lehrkräfte. Opladen und Farmington Hills: Barbara Budrich.

Joughin, G. (2010). A short guide to oral assessment. Leeds: Metropolitan University / University of Wollongong.

Nallaya, S., Gentili, S., Weeks, S. & Baldock, K. (2024). The validity, reliability, academic integrity and integration of oral assessments in higher education: A systematic review. Issues in Educational Research, 34(2).

Mattes, W. (2011). Methoden für den Unterricht. Kompakte Übersichten für Lehrende und Lernende. Braunschweig: Westermann.

Looney, J. (2011). Integrating Formative and Summative Assessment: Progress Toward a Seamless System? OECD Education Working Papers No. 58. Paris: OECD.

Frenzel, A., Götz, T., Pekrun, R. (2015). Emotionen. In: Pädagogische Psychologie. 2. Aufl. Hg. v. Wild, E., Möller, J. Heidelberg: Springer.

Götz, T., Zirngibl, A., Pekrun, R. (2004). Lern- und Leistungseemotionen von Schülerinnen und Schülern. In: Schule positive erleben. Ergebnisse und Erkenntnisse zum Wohlbefinden von Schülerinnen und Schülern. Schulpädagogik – Fachdidaktik – Lehrerbildung. Hg. v. Hascher, T.

Metha, B., Bhandari, B., Sharma, P. & Kaur, R. (2016). Short Answer Open-Ended versus Multiple-Choice Questions: A Comparison of Objectivity.

## Bearbeitungsvermerk

Dieser Beitrag wurde inhaltlich und illustrativ mit Unterstützung von künstlicher Intelligenz (KI) erstellt. Er wurde von den Autorinnen und Autoren fachlich geprüft, angepasst und freigegeben.