

5 Wirksamkeitsevaluation

Übungsaufgaben

- (1) Eine Organisationspsychologin möchte untersuchen wie sich eine erweiterte Home-Office Regelung auf die Arbeitsmotivation der Angestellten auswirkt. Dazu befragt sie die Angestellten, die entschieden haben vorwiegend im Home-Office zu arbeiten, nach ihrer Arbeitsmotivation und vergleicht diese mit den Antworten von Angestellten, die entschieden haben im Büro zu arbeiten.
- (a) Erläutern Sie anhand dieses Beispiels was maßnahmenspezifische Wirkungen, Neben- und Folgewirkungen, maßnahmenunspezifische Wirkungen und externe Wirkungen sind.

Musterantwort

Maßnahmenspezifische Wirkungen sind intendierte kausale Wirkungen, die auf den Hauptwirkstoff zurückzuführen sind. In unserem Beispiel wäre es der kausale Effekt des Arbeitsortes (hauptsächlich Homeoffice vs. Büro) auf die Arbeitsmotivation. Neben- und Folgewirkungen sind ebenfalls maßnahmenspezifisch, jedoch nicht intendiert. In dem Beispiel könnten Neben- und Folgewirkungen von Homeoffice im Vgl. zu Büro als Arbeitsort eine verringerte Kommunikation im Team oder ein geringerer Teamzusammenhalt sein. Maßnahmenunspezifische Wirkungen sind nicht auf den Hauptwirkstoff der Maßnahme, sondern auf andere Bestandteile der Intervention zurückzuführen. So könnte es im vorliegenden Beispiel sein, dass Angestellte im Büro nun mehr Aufgaben übernehmen müssen, die vor Ort anfallen als Angestellte im Home-Office. Unterschiede in der Aufgabenverteilung gehören nicht zum ‚Hauptwirkstoff‘ des Vergleiches Home-Office vs. Büro und zählt daher zu den maßnahmenunspezifischen Wirkungen. Externe Wirkungen gehen auf systematische Verzerrungen zurück. In unserem Beispiel könnte es sein, dass insbesondere Personen mit geringer Arbeitsmotivation sich entscheiden im Homeoffice zu arbeiten. Unterschiede in der Arbeitsmotivation zwischen Personen im Homeoffice und im Büro sind dann nicht notwendigerweise auf den Arbeitsort, sondern auf bereits zuvor bestehende Unterschiede zurückzuführen.

- (b) Geben Sie an, welche der in (a) genannten Wirkungen zur Brutto-Wirkung und welche zur Netto-Wirkung gezählt werden.

Musterantwort

Alle vier Wirkungen zusammen werden als Brutto-Wirkung bezeichnet. Sie sind die Wirkung, die wir beobachten. Die maßnahmenspezifischen Wirkungen und Neben- sowie Folgewirkungen bezeichnet man als Netto-Wirkung. Diese beschreiben den kausalen Effekt des ‚Hauptwirkstoffs‘.

- (2) Sie möchten den Effekt von interaktiven Elementen im Unterricht (z.B. Kleingruppenarbeit oder Diskussionen) auf die Lernmotivation untersuchen. Erläutern Sie, welche Art von Kontrollgruppe Sie wählen würden. Was sollten Sie bei der Wahl der Kontrollgruppe beachten?

Musterantwort

Da das Ziel ist, speziell den Effekt von interaktiven Elementen im Unterricht zu untersuchen, nicht jedoch den gesamten Unterricht, sollte eine unspezifisch behandelte Kontrollgruppe gewählt werden. Diese sollte sich idealerweise von der Interventionsgruppe nur in der Verwendung interaktiver Elemente, nicht jedoch in anderen Aspekten unterscheiden. So könnte man den Unterricht in der Kontrollgruppe zum Beispiel mit dem gleichen Inhalt, von der gleichen Lehrkraft und zu einer ähnlichen Uhrzeit wie den Unterricht in der Interventionsgruppe durchführen, und nur die interaktiven Elemente durch Frontalunterricht ersetzen.

- (3) Sie untersuchen die Wirkung eines dreimonatigen Strategietrainings zur Erhöhung der Merkfähigkeit bei Grundschulkindern. Dazu erfassen Sie die Merkfähigkeit der Kinder vor dem Training sowie nach dem Training und stellen fest, dass sich die Merkfähigkeit erhöht hat.
- (a) Warum kann von diesem Ergebnis keine sichere Schlussfolgerung auf die Wirkung des Trainings gezogen werden?

Musterantwort

Dadurch, dass nur eine Gruppe und diese nur einmal vor und einmal nach dem Training untersucht wird, sind viele Alternativerklärungen für die Erhöhung der Merkfähigkeit denkbar. Dazu gehören:

- *Externe zeitliche Einflüsse – z.B. könnte es sein, dass die Kinder auch im regulären Unterricht Lernstrategien vermittelt bekommen haben. Sie könnten somit nicht nur durch das Strategietraining ihre Merkfähigkeit verbessert haben, sondern auch durch den Unterricht.*
- *Reifung – es ist plausibel, dass Kinder im Laufe der Zeit auch ohne Training, zum Beispiel durch die fortschreitende Entwicklung ihres Gehirns, ihre Merkfähigkeit verbessern.*
- *Testübung – die Kinder könnten sich an Teile des ersten Tests zur Merkfähigkeit erinnern oder beim zweiten Test bereits vertrauter mit dem Test sein. Eine höhere Merkfähigkeit nach dem Training kann dann auch aufgrund dieser Effekte zurückgehen und muss nicht (nur) durch das Training erfolgt sein.*
- *eine Veränderung im Messinstrument – wenn beispielsweise beim Test zur Merkfähigkeit nach dem Training schwierigere Aufgaben als im Test vor dem Training verwendet werden und für diese Schwierigkeitsunterschiede nicht kontrolliert wird, kann ein Unterschied zwischen den Testergebnissen vorher und nachher auf Schwierigkeitsunterschiede zurückgehen und muss eben nicht (nur) den Effekt des Trainings beschreiben.*
- *Regression zur Mitte – Regression zur Mitte tritt auf, wenn die Merkfähigkeitswerte zum zweiten Messzeitpunkt allein aufgrund eines statistischen Artefakts weniger extrem sind als zum ersten Messzeitpunkt, insbesondere wenn die Merkfähigkeit der Kinder zum ersten Messzeitpunkt besonders extrem war, z.B. besonders schlecht. Eine Verbesserung der Merkfähigkeit muss dann nicht nur aufgrund des Trainings, sondern kann auch aufgrund des statistischen Artefaktes erfolgt sein.*
- *Selektiver Drop-out – Wenn beispielsweise Kinder mit einer besonders guten Merkfähigkeit das Training abgebrochen haben, weil sie unterfordert waren und entsprechend an der Testung nach dem Training fehlen, kann das die mittlere Merkfähigkeit nach dem Training verringern. Würde man den Unterschied als kausalen Effekt des Trainings interpretieren, würden wir fälschlicherweise den Effekt des Trainings unterschätzen.*

- (b) Wie könnte man eine kausale Wirkung des Trainings empirisch stärker absichern?

Musterantwort

Durch die Hinzunahme einer Kontrollgruppe könnten die in a) genannten Gefahren (Veränderungen über die Zeit) kontrolliert werden. Durch die Hinzunahme einer Kontrollgruppe könnten nun jedoch Selektionseffekte oder Interaktionen von Selektion und Veränderung über die Zeit auftreten. Um diese zu kontrollieren kann eine Randomisierung der Personen zu den Interventionsgruppen erfolgen. Ist eine Randomisierung nicht möglich, können Störvariablen vor der Zuweisung zu den Interventionsbedingungen erhoben und per Design oder statistisch kontrolliert werden.

Statt einer Kontrollgruppe kann man auch die Anzahl der Messzeitpunkte vor und nach der Intervention erhöhen, beispielsweise die Merkfähigkeit monatlich zu vier Zeitpunkten vor dem Training und viermal nach dem Training messen. Man kann so Veränderungen, die unabhängig von Training auftreten, beobachten und ‚verrechnen‘. Annahme ist jedoch, dass die Gefahren, die während des Trainings auftreten aufgrund der

Veränderung vor dem Training abgeschätzt werden können. Das ist bei externen zeitlichen Einflüssen meist etwas schwieriger anzunehmen, jedoch für die anderen Gefahren über die Zeit meist plausibel.

- (4) Frau Dr. A. Rückriem hat eine klinische Intervention zur Behandlung von sozialen Ängsten entwickelt, in der Methoden aus verschiedenen Therapieschulen zum Einsatz kommen. Sie möchte herausfinden, welchen Effekt ihre Therapie auf die Lebenszufriedenheit aller Teilnehmenden hat. Für ihre Untersuchung plant Frau Dr. Rückriem, alle an der Studie Interessierten randomisiert entweder der Therapie- oder einer Kontrollgruppe zuzuweisen. Sie will die AV bei allen Studienteilnehmenden einmal vor der Zuweisung zu den Gruppen und einmal nach Beendigung der Therapie messen.
- (a) Die Ethikkommission hat Bedenken, dass insbesondere den hoch belasteten Patient:innen keine Behandlung angeboten wird. Nennen Sie zwei verschiedene Möglichkeiten, die ethischen Bedenken aufzugreifen und dennoch eine Form der Randomisierung umzusetzen.

Musterantwort

Möglichkeit 1: Randomisierung nach Einwilligung – Es werden nur Patient:innen in die Studie eingeschlossen, die einwilligen (oder bei denen die/der Therapeut:in zustimmt) sich randomisiert einer Behandlung oder Kontrollbedingung zuweisen zu lassen. Patient:innen, die so stark belastet sind, dass sie sofort eine Behandlung benötigen, würden in der Regel nicht in die Studie aufgenommen. Mit allen eingewilligten Personen kann nun eine Randomisierung durchgeführt werden. Jedoch ist die Generalisierbarkeit der Ergebnisse (externe Validität) eingeschränkt, da nur eine selektive Auswahl aus der Population untersucht wurde.

Möglichkeit 2: Randomisierung mit Warte-Kontrollgruppe – Wenn es ethisch vertretbar ist, dass die Patient:innen nicht sofort, sondern nur etwas später eine Behandlung erhalten, kann auch eine Warte-Kontrollgruppe in Betracht gezogen werden. Hier würden Patient:innen der Kontrollbedingung nach Ende der Studie die Behandlung erhalten.

Möglichkeit 3: Kontrollgruppe mit Alternativbehandlung – Statt einer unbehandelten Kontrollgruppe, kann eine Kontrollbedingung mit einer alternativen Behandlung, z.B. einer herkömmlichen Therapie oder eines Achtsamkeitstrainings, angeboten werden. Dies stellt sicher, dass Personen in der Kontrollbedingung nicht ohne Behandlung bleiben. Allerdings ändert sich damit die Definition der UV. Es wird nicht der Gesamteffekt der neuen Intervention untersucht, sondern der Unterschied der neuen Intervention im Vergleich zu den bisherigen Behandlungen.

Möglichkeit 4: Bedingte Randomisierung – Die Wahrscheinlichkeit in die Interventionsgruppe zu gelangen wird für stark belastete Patient:innen deutlich höher als für weniger belastete Patient:innen gesetzt; beispielsweise werden stark belastete Patient:innen mit 75% Wahrscheinlichkeit der Therapie zugewiesen, während nicht stark belastete Patient:innen nur mit 30% Wahrscheinlichkeit die Therapie erhalten. Damit wird sichergestellt, dass ein Großteil der stark belasteten Patient:innen die Therapie erhält. Die unterschiedlichen Wahrscheinlichkeiten in die Therapie- bzw. Kontrollgruppe zu gelangen, werden dann in der Analyse berücksichtigt.

Möglichkeit 5: Regressions-Diskontinuitäts-Design – Alle Personen, die einen bestimmten Grenzwert der der Belastung überschreiten, erhalten die Therapie, die anderen werden der Kontrollbedingung zugeordnet. Somit wird sichergestellt, dass alle stark belasteten Patient:innen die Therapie erhalten. In der Analyse ist in diesem Fall allerdings lediglich die Schätzung des kausalen Effekts der Behandlung am Grenzwert möglich.

- (b) Frau Dr. A. Rückriem entscheidet sich nach längeren Überlegungen dazu, auf die Randomisierung zu verzichten und die Studienteilnehmenden selbst entscheiden zu lassen, ob sie an der Intervention teilnehmen möchten oder nicht. Welche Gefahren für die interne Validität treten in diesem Design auf? Nennen Sie zwei Möglichkeiten, wie Sie diese Gefahren möglichst gering halten können.

Musterantwort

In diesem Design besteht die Gefahr darin, dass Selektionseffekte sowie Interaktionen von Selektionseffekten und Veränderungen über die Zeit auftreten können. Das heißt, dass es passieren kann, dass Gruppenunterschiede in der Lebenszufriedenheit (auch) durch Störvariablen statt (allein) durch die Therapie bedingt sein können. Ebenso können Gruppenunterschiede in der Lebenszufriedenheit durch die Interaktion von Störvariablen und Veränderungen über die Zeit entstehen.

Um diese Gefahren möglichst gering zu halten, können vor der Durchführung der Studie möglichst viele relevante Störvariablen identifiziert und in der Studie erfasst werden. Relevant sind hierbei alle möglichen Variablen, anhand derer sich Gruppen unterscheiden könnten (Zusammenhang mit Interventionsvariable) und die voraussichtlich mit der Lebenszufriedenheit zusammenhängen (Zusammenhang mit Outcome). Diese Störvariablen können dann entweder per Design oder statistisch konstant gehalten werden.

Möglichkeit 1: Konstanthaltung – Besonders relevante Störvariablen (z.B. die Schwere der Symptomatik) können konstantgehalten werden, indem beispielsweise nur schwer belastete Patient:innen an der Studie teilnehmen. Dies schränkt jedoch die externe Validität der Studie ein, da durch dieses Vorgehen Aussagen über den Effekt der Therapie nur noch in Bezug auf diese Gruppe möglich sind.

Möglichkeit 2: Ausbalancierung – Bei der Ausbalancierung werden Störvariablen in das Design mit aufgenommen. Beispielsweise kann die Symptombelastung vorher als weiterer Faktor mit aufgenommen werden. Personen werden so auf die vier Bedingungen verteilt, dass die Korrelation zwischen Behandlungsgruppe und Symptombelastung null ist; beispielsweise 100 Personen in jeder der vier Gruppen. Somit wird sichergestellt, dass die Verteilung der Störvariablen in beiden Gruppen der UV gleich ist.

	Geringe Symptombelastung	Hohe Symptombelastung
Therapie	n=100	n=100
Kontrolle	n=100	n=100

Möglichkeit 3: Parallelisierung – Bei der Parallelisierung wird für jede Person in der Therapiegruppe eine ‚passende‘ Person in der Kontrollgruppe gewählt, die in den Störvariablen ähnlich ist. Diese nennt man auch ‚statistische Zwillinge‘. So wird beispielsweise für einen 30-jährigen Patienten mit hoher Symptombelastung in der Therapiegruppe eine Person in ähnlichem Alter und mit ähnlicher Symptombelastung für die Kontrollgruppe gewählt. Damit wird sichergestellt, dass die Gruppen sich in der Verteilung der betrachteten Störvariablen ähnlich sind.

Möglichkeit 4: statistische Kontrolle – Der Einfluss der Störvariablen kann mit statistischen Verfahren, wie beispielsweise der Kovarianzanalyse, Kovariaten-Matching oder Propensity-Score-Verfahren, kontrolliert werden.