

**Bericht zur Evaluation
der ersten Trainingswelle des Gewaltpräventionsprojekts
„PiT-Hessen (Prävention im Team)“**

**im Auftrag der
PiT-Lenkungsgruppe**



Dipl. Psych. Gunnar Lemmer

Dipl. Psych. Claudia Neumann, Prof. Dr. Ulrich Wagner

Fachbereich Psychologie
der Philipps-Universität Marburg
AG Sozialpsychologie

Dezember 2005

Kontakt:

Fachbereich Psychologie
der Philipps-Universität Marburg
AG Sozialpsychologie

Briefe: 35032 Marburg
Pakete: Gutenbergstraße 18, 35037 Marburg

<http://www.sozialpsychologie-marburg.de>

Dipl.-Psych. Gunnar Lemmer,¹

Telefon: 06421/28-23654
Telefax: 06421/28-23789

e-mail: g.lemmer@staff.uni-marburg.de

Prof. Dr. Ulrich Wagner,

Telefon: 06421/28-23664/8
Telefax: 06421/28-23789

e-mail: Wagner1@staff.uni-marburg.de

¹ Wir Danken Judith Lanphen, Ingela Mänz, Julia Otto, Katharina Plutta, Uli Frischknecht u. Silke Eschert für die Unterstützung bei der Dateneingabe u. -aufbereitung.

Inhaltsverzeichnis

1. Zusammenfassung	5
2. Hintergrundinformationen	7
2.1. Evaluationsauftrag	8
2.2. Das Projekt „PiT-Hessen“	8
2.2.1. Projekthintergrund	8
2.2.2. Angaben zu den Trainingsmaßnahmen	9
2.2.2.1. Trainingsziele	9
2.2.2.2. Aufbau des Trainings	10
2.2.2.3. Durchführung des Trainings – das Modellprojekt Frankfurt-Offenbach	11
2.3. Allgemeine Bemerkungen zur Projektevaluation	13
2.3.1. Evaluation und Projektevaluation	13
2.3.2. Ergebnisevaluation	14
2.3.2.1. Trainingsziele	14
2.3.2.2. Messinstrument	14
2.3.2.3. Untersuchungsdesign	15
2.3.3. Prozessevaluation	17
3. Ergebnisevaluation	18
3.1. Methodik	19
3.1.1. Messinstrument	19
3.1.1.1. Allgemeines zum Messinstrument	19
3.1.1.2. Vorstellung der einzelnen Indikatoren	22
3.1.2. Untersuchungsdesign	27
3.2. Auswertungsergebnisse	29
3.2.1. Angaben zur Analysestichprobe	29
3.2.1.1. Voraussetzungen zur Aufnahme in die Analysestichprobe	29
3.2.1.2. Kennzeichnung der Analysestichprobe	30
3.2.1.3. Soziodemografische Beschreibung der Analysestichprobe	31
3.2.2. Empirische Skalenkonstruktion – Struktur und Qualität der Indikatoren bzw. Skalen	32
3.2.2.1. Einführende Bemerkungen	32
3.2.2.2. Ergebnis der Faktorenanalysen – Struktur der Indikatoren	33
3.2.2.3. Reliabilitätsanalysen – Qualität der Skalen	35
3.2.2.4. Fazit zur empirischen Skalenkonstruktion	36
3.2.3. Effektivität des Trainings	37
3.2.3.1. Einführende Bemerkungen	37
3.2.3.2. Effektivität des Trainings I – Ergebnis der multivariaten Analyse	41
3.2.3.3. Effektivität des Trainings II – Ergebnis der univariaten Analysen	42
3.2.3.4. Effektivität des Trainings III – zum Einfluss weiterer Faktoren	48
3.3. Fazit zur Ergebnisevaluation	50

4. Prozessevaluation	51
4.1. Prozessevaluation – TeilnehmerInnen	52
4.1.1. Methodik: Messinstrument u. Untersuchungsdesign	52
4.1.2. Auswertungsergebnisse	54
4.1.2.1. Analysestichprobe	54
4.1.2.2. Auswertungsergebnisse – geschlossene Fragen	54
4.1.2.3. Auswertungsergebnisse – offene Fragen	56
4.2. Prozessevaluation – TeamarbeiterInnen	59
4.2.1. Methodik: Messinstrument u. Untersuchungsdesign	59
4.2.2. Auswertungsergebnisse	60
4.2.2.1. Auswertungsergebnisse – geschlossene Fragen	60
4.2.2.2. Auswertungsergebnisse – offene Fragen	62
4.3. Fazit zur Prozessevaluation	64
5. Diskussion	65
6. Literaturverzeichnis	67
7. Anhang	68

1. Zusammenfassung

Der folgende Bericht enthält den Befund einer von der Arbeitsgruppe Sozialpsychologie der Philipps-Universität Marburg durchgeführten wissenschaftlichen Ergebnis- und Prozessevaluation der ersten Welle der im Rahmen des Gewaltpräventionsprogramms „PiT-Hessen (Prävention im Team)“ durchgeführten Trainingsmaßnahmen.

Das Projekt PiT-Hessen basiert auf einer Initiative der Hessischen Landesregierung im Rahmen des „Netzwerks gegen Gewalt“ und wird vom Hessischen Ministerium des Innern und für Sport, dem Hessischen Kultusministerium, dem Hessischen Sozialministerium und dem Hessischen Justizministerium getragen.

Eine Besonderheit von PiT-Hessen ist, dass es im Vergleich zu anderen Präventionsprogrammen explizit auf der Kooperation von Schule, Polizei und Jugendhilfe beruht. Das im Rahmen des Projekts durchgeführte Training besteht insgesamt aus vier Bausteinen (Was ist Gewalt, Konflikteinstieg, Konfliktausstieg und Helferverhalten), denen jeweils zwischen acht und 14 Übungen zugeordnet sind.

Ziel des Trainings ist es, Jugendliche zum Erkennen von Konfliktsituationen und zu einem an den persönlichen Möglichkeiten orientierten Umgang mit Konfliktsituationen zu befähigen. Zielgruppe sind dabei insbesondere Schülerinnen und Schüler der siebten bis achten Klassen der allgemein bildenden Schulen.

Seit dem Schuljahr 2004/2005 wird an 16 Schulen in Frankfurt am Main sowie der Stadt und dem Kreis Offenbach in einem Modellprojekt die Praxistauglichkeit von PiT-Hessen erprobt. An jeder der beteiligten Projektschulen wurde ein Team aus mindestens zwei LehrerInnen, einer Polizeibeamtin oder einem Polizeibeamten sowie einer Sozialarbeiterin oder einem Sozialarbeiter aus Einrichtungen der Jugendhilfe etabliert. Die Teammitglieder sind an den Schulen kooperativ als Trainer tätig. Im Rahmen dieser Tätigkeit übernehmen sie für die betreffende Schule gemeinsam die Verantwortung für Inhalt, Umfang und Durchführung der durchgeführten Trainingsmaßnahmen.

Die Arbeitsgruppe Sozialpsychologie der Universität Marburg wurde im Juli 2004 von der PiT-Lenkungsgruppe beauftragt, die im Rahmen des Modellprojekts durchgeführte erste Trainingswelle wissenschaftlich zu evaluieren. Die Evaluation wurde von der AG Sozialpsychologie in Form einer Ergebnisevaluation sowie einer Prozessevaluation angelegt und durchgeführt.

Im Rahmen der Ergebnisevaluation wurde überprüft, ob durch das Training die Trainingsziele erreicht werden. Das bereits geschilderte Trainingsziel wurde dazu von den InitiatorInnen in neun Teilziele ausdifferenziert (z.B. Entwicklung persönlicher Handlungsoptionen). Zur Ermöglichung der Ergebnisevaluation wurde von der AG Sozialpsychologie ein Fragebogen konstruiert. Dieses Messinstrument enthält Indikatoren für die Ausprägung der in den Teilzielen thematisierten Fähigkeiten bzw. Merkmale (z.B. Fähigkeit zur Vermeidung eines Konflikteinstiegs). Jedem Indikator sind mehrere Items (Aussagen) zugeordnet, zu denen die SchülerInnen im Rahmen der Befragung anhand vorgegebener Antwortkategorien Stellung nehmen sollten. Die Indikatoren sind auf vier verschiedenen Ebenen angesiedelt: Verhaltensintention, Situationsbewertung, Einstellung und retrospektive Verhaltensangabe.

Um ermitteln zu können, ob es in den Indikatoren durch das Training zu einer zielgemäßen Veränderung kommt, wurde im Rahmen eines Vortest-Nachtest-Untersuchungsdesigns mit nicht-äquivalenter Kontrollgruppe eine wiederholte Befragung von Trainings- und Kontrollklassen vor und nach der Durchführung der Trainingsmaßnahmen durchgeführt.

In der Auswertung der erhobenen Daten wurde zunächst die Analysestichprobe aus SchülerInnen, die den Fragebogen zu beiden Messzeitpunkten ausgefüllt haben und bei denen eine Verknüpfung der beiden Dokumente möglich war, zusammengestellt. Anschließend wurden aus den Indikatoren durch Faktoren- und Reliabilitätsanalysen elf empirisch abgesicherte Skalen (Inhaltsbereiche) konstruiert (z.B. Konflikteinstieg).

Zur Analyse der Trainingseffektivität wurden inferenzstatistische Verfahren (multi- und univariate Varianzanalysen mit Messwiederholung) herangezogen, mit denen die Befragungswerte vor und nach Durchführung der Trainingsmaßnahmen getrennt für Trainingsgruppe und Kontrollgruppe verglichen wurden. Durch diese Verfahren lässt sich zum Beispiel ermitteln, ob im Sinne der Ziele in der Trainingsgruppe - im Gegensatz zur Kontrollgruppe - über die Zeit hinweg eine positive Veränderung stattfindet und, ob mit hinreichender Sicherheit davon auszugehen ist, dass „in Wahrheit“ ein Trainingseffekt vorhanden ist oder, ob wahrscheinlich ist, dass die Veränderung lediglich auf zufällige Variationen der herangezogenen Stichprobe zurückzuführen ist. Die Ergebnisse zeigen in vier der elf Skalen einen Trainingseffekt (Konflikteinstieg, Konfliktausstieg, Zivilcourage 1 - Kommunikation mit Umstehenden, Zivilcourage 2 - direktes Eingreifen): Die Merkmalsausprägung steigt in der Trainingsgruppe und sinkt in der Kontrollgruppe. Die Variablen Geschlecht, Schultyp, Jahrgangsstufe und Migrationshintergrund haben keinen bedeutsamen Einfluss auf die ermittelte Trainingseffektivität.

Die aufgeführten Skalen sind auf der Ebene der Verhaltensintention lokalisiert. Gemäß den Ergebnissen ist daher davon auszugehen, dass die TrainingsteilnehmerInnen durch das Training primär Skripte (Verhaltensrichtlinien) erlernen, die beinhalten, wie sie sich in konkreten Situationen mit einem gewissen Konfliktpotential zu verhalten haben.

Im Rahmen der Prozessevaluation wurde untersucht, wie das Training aus der Perspektive der TrainingsteilnehmerInnen (Trainingsgruppe) und aus der Perspektive der TeamarbeiterInnen bewertet wird. Für beide Gruppen wurde ein geeignetes Messinstrument mit mehreren geschlossenen und offenen Fragen konstruiert.

Die TrainingsteilnehmerInnen wurden mit acht geschlossenen Fragen und zwei offenen Fragen befragt. Die Auswertung zeigt, dass die Trainingsteilnehmer das Training insgesamt positiv bewerten. Den TeamarbeiterInnen wurden 15 geschlossene und drei offene Fragen vorgelegt. Auch aus dieser Perspektive ergibt sich ein insgesamt positives Bewertungsbild.

Vor dem Hintergrund dieser Ergebnisse ist die weitere Verbreitung und Umsetzung des Trainings aus wissenschaftlicher Sicht zu empfehlen.

2. Hintergrundinformationen

In diesem Kapitel wird auf den Hintergrund der geschilderten Projektevaluation eingegangen. Das Kapitel ist in folgende Unterkapitel gegliedert:

- 2.1. Evaluationsauftrag.
- 2.2. Das Projekt „PiT-Hessen“.
- 2.3. Allgemeine Bemerkungen zur Projektevaluation.

Zunächst wird kurz der Evaluationsauftrag skizziert. Anschließend werden einige zentrale Aspekte des Projekts PiT-Hessen vorgestellt. Zum Abschluss wird auf einige allgemeine Aspekte der Projektevaluation eingegangen.

2.1. Evaluationsauftrag

Die Arbeitsgruppe Sozialpsychologie der Universität Marburg wurde im Juli 2004 von der PiT-Lenkungsgruppe beauftragt, die im Rahmen des PiT-Modellprojekts Frankfurt-Offenbach im Schuljahr 2004 / 2005 durchgeführte erste Trainingswelle wissenschaftlich zu evaluieren. Die Projektevaluation wurde von der AG Sozialpsychologie in Absprache mit den Auftraggebern in Form einer Ergebnis- und Prozessevaluation konzipiert und durchgeführt.

2.2. Das Projekt „PiT-Hessen“

2.2.1. Projekthintergrund

Das Projekt basiert auf einer Initiative der Hessischen Landesregierung im Rahmen des „Netzwerks gegen Gewalt“ und wird vom Hessischen Ministerium des Inneren und für Sport, dem Hessischen Kultusministerium, dem Hessischen Sozialministerium und dem Hessischen Justizministerium getragen.

PiT-Hessen ist ein Gewaltpräventionsprogramm. Es verfolgt nach Angabe der InitiatorInnen drei Projektziele:

1. Teambildung aus Personen von Schule, Polizei und Jugendhilfe.
2. Trainings mit Schülerinnen und Schülern.
3. Impulse für die Personal-, Organisations- und Konzeptentwicklung insbesondere in der Schule geben.

Eine Besonderheit von PiT-Hessen ist der zentrale Stellenwert einer Teambildung aus Mitgliedern verschiedener Berufsgruppen. Das Konzept der Teambildung beruht dabei auf der Kooperation der drei Berufsgruppen Schule, Polizei und Jugendhilfe bei der Planung und Durchführung der Trainingsmaßnahmen (siehe später). Ein derartiger Teambildungs-Aspekt wird von vergleichbaren Präventionsprogrammen in der Regel nicht oder in deutlich geringerem Ausmaß berücksichtigt.

Im Rahmen dieses Berichts wurden in erster Linie die Trainingsmaßnahmen evaluiert. Aus den Befunden lassen sich jedoch auch Aussagen zur Evaluation der Teambildung ableiten. Die Evaluation des Projektziels „Impulse für die Personal-, Organisations- und Konzeptentwicklung insbesondere in der Schule geben“ ist Bestandteil einer weiteren Evaluation. Diesbezüglich ist geplant, in einigen Projektschulen Interviews zur Ausstrahlungswirkung der dort mit den Trainingsklassen durchgeführten Maßnahmen durchzuführen und inhaltsanalytisch auszuwerten.

2.2.2. Angaben zu den Trainingsmaßnahmen

In diesem Abschnitt werden die folgenden Unterpunkte behandelt:

- 2.2.2.1. Trainingziele.
- 2.2.2.2. Aufbau des Trainings.
- 2.2.2.3. Durchführung des Trainings – das Modellprojekt Frankfurt-Offenbach.

Zunächst werden die Trainingsziele geschildert. Als nächstes wird kurz der Aufbau des Trainings beschrieben, bevor dann auf die Durchführung des Trainings im Rahmen des Modellprojekts Frankfurt-Offenbach eingegangen wird.

2.2.2.1. Trainingziele

Das allgemeine Ziel der im Rahmen des Gewaltpräventionsprogramms PiT-Hessen durchgeführten Trainingsmaßnahmen besteht nach Aussage der InitiatorInnen darin, Jugendliche zum Erkennen von Konfliktsituationen und zu einem an den persönlichen Möglichkeiten orientierten Umgang mit Konfliktsituationen zu befähigen. Das Training möchte diesbezüglich mit einem gewaltfreien Ansatz Optionen aufzeigen und den Teilnehmern die Möglichkeit bieten, an den persönlichen Möglichkeiten des Einzelnen orientierte Handlungsoptionen zur gewaltfreien Lösung von Konfliktsituationen zu entwickeln.

Zielgruppe des Trainings sind nach Angabe der InitiatorInnen insbesondere Schülerinnen und Schüler der siebten bis achten Klassen allgemein bildender Schulen.

Im Vorfeld der Projektevaluation wurde das oben genannte allgemeine Ziel von den InitiatorInnen des Trainings in neun konkrete Teilziele ausdifferenziert. Diese Teilziele sind für eine Ergebnisevaluation gut geeignet (siehe später). Jedes Ziel ist als positive Veränderung des Ausmaßes einer Fähigkeit zum Erkennen von Konfliktsituationen oder zum Umgang mit Konfliktsituationen formuliert (z.B. Stärkung von Zivilcourage). Die Teilziele werden in Tabelle 1 aufgeführt.

Tabelle 1

Teilziele der PiT-Trainingsmaßnahmen

1. Entwicklung persönlicher Handlungsoptionen.
2. Hinsehen statt Wegesehen.
3. Stärkung von Zivilcourage.
4. Erlernen von „richtigem“ Helferverhalten in gewaltbesetzten Situationen.
5. Frühzeitiges Einbinden von „Helfern“ in Problemsituationen.
6. Sensibilisierung für das frühzeitige Erkennen konflikthafter bzw. gefahrenträchtiger Situationen.
7. Weg von der passiven Opferrolle hin zum reflektierten Umgang mit gewaltbesetzten Situationen.
8. Vertrauensbildung/-Intensivierung gegenüber LehrerInnen in deren Rolle als Hilfsperson.
9. Reflexion der eigenen Ängste und Befürchtungen.

2.2.2.2. Aufbau des Trainings

Das Training basiert auf dem Reader „PiT - Prävention im Team. Ein hessisches Gewaltpräventionsprogramm mit Teambildung von Schule, Polizei und Jugendhilfe“. Ein Teil der dort aufgeführten Trainingsmaßnahmen wurde eigens für PiT-Hessen entwickelt, ein anderer Teil entstammt dem vom Frankfurter Kinderbüro entwickelten Trainingsprogramm „Cool sein – cool bleiben“.

Das Training umfasst laut Trainingsreader insgesamt vier Bausteine:

1. Was ist Gewalt.
2. Konflikteinstieg.
3. Konfliktausstieg.
4. Helferverhalten.

Die Bausteine beziehen sich jeweils auf einen bestimmten thematischen Schwerpunkt und werden im Verlauf des Trainingsprozesses nacheinander bearbeitet. Jedem Baustein sind zwischen acht und 14 Übungen zugeordnet. In den Übungen wird sich in unterschiedlicher Weise mit dem jeweiligen thematischen Schwerpunkt des Bausteins auseinandergesetzt. Die Anzahl und Auswahl der pro Baustein durchzuführenden Übungen ist nicht festgesetzt, die Übungen können auch alternativ eingesetzt werden. Darüber hinaus besteht keine fixe Vorgabe hinsichtlich der Art der Durchführung des Trainings (z.B. im Rahmen einer Projektwoche vs. im Rahmen einzelner Schulstunden). Die Entscheidungen zu diesen Aspekten unterliegen der Verantwortung des an der betreffenden Schule tätigen Teams.

2.2.2.3. Durchführung des Trainings – das Modellprojekt Frankfurt-Offenbach

Seit dem Schuljahr 2004/2005 wird PiT-Hessen hinsichtlich seiner Praxistauglichkeit in einem Modellprojekt getestet. Dieses Modellprojekt wird in insgesamt 16 Schulen in Frankfurt am Main sowie der Stadt und dem Kreis Offenbach durchgeführt. Es handelt sich dabei um vier Gymnasien, zwei Realschulen, zwei Haupt- und Realschulen, zwei Hauptschulen sowie um sechs Gesamtschulen. Die Schulen sind in Tabelle 2 aufgeführt.

Tabelle 2

Schule	Schultyp	Ort
1. Albert-Einstein-Schule	Gymnasium	Langen
2. Albert-Schweitzer-Schule	Gesamtschule	Offenbach am Main
3. Bornheimer Realschule	Realschule	Frankfurt am Main
4. Carlo-Mierendorff-Schule	Gesamtschule	Frankfurt am Main
5. Edith-Stein-Schule	Gesamtschule	Offenbach am Main
6. Fürstenbergerschule	Realschule	Frankfurt am Main
7. Geschw.-Scholl-Schule	Haupt- u. Realschule	Offenbach am Main
8. Helmholtzschule	Gymnasium	Frankfurt am Main
9. Hostatoschule	Hauptschule	Frankfurt am Main
10. IGS Nordend	Gesamtschule	Frankfurt am Main
11. Mathildenschule	Haupt- u. Realschule	Offenbach am Main
12. Oswald-v-Nell-Breuning-Schule	Gesamtschule	Rödermark
13. Otto-Hahn-Schule	Gesamtschule	Frankfurt am Main
14. Rudolf-Koch-Schule	Gymnasium	Offenbach am Main
15. Schillerschule	Gymnasium	Frankfurt am Main
16. Schwanthalerschule	Hauptschule	Frankfurt am Main

Die erste Trainingswelle wurde gegen Ende des Schuljahres 2004 / 2005 abgeschlossen. Die oben genannten Schulen nahmen an dieser Trainingswelle mit jeweils zwischen einer und fünf Trainingsklassen sowie zwischen einer und zwei Kontrollklassen teil (näheres siehe später). Die SchülerInnen der Kontrollklassen erhielten kein Training. Mit den Trainingsklassen wurden in jeder Schule die vier thematischen Bausteine des Trainings behandelt.

Gemäß des Grundkonzepts von PiT-Hessen lag die Verantwortung hinsichtlich Inhalt, Ausmaß und Durchführung der einzelnen Übungen sowie hinsichtlich der generellen Durchführungsform des Trainings (z.B. in einzelnen Projekttagen vs. im Stundenplan integriert) innerhalb der Schulen der Verantwortlichkeit eines dort gebildeten Trainerteams aus mindestens zwei LehrerInnen, einer Polizeibeamtin oder einem Polizeibeamten sowie einer Sozialarbeiterin oder einem Sozialarbeiter aus Einrichtungen der Jugendhilfe. Aus diesem Grund variierte sowohl die Art und Anzahl der im Rahmen der vier Bausteine durchgeführten Übungen als auch die Art der Durchführung des Trainings zwischen den Schulen. Die Trainings sind daher hinsichtlich dieser Aspekte sehr heterogen.

2.3. Allgemeine Bemerkungen zur Projektevaluation

Der Begriff „Evaluation“ beinhaltet eine Vielzahl von Bedeutungsfacetten und Vorgehensweisen. In der Literatur finden sich zahlreiche, zum Teil sehr divergierende Vorstellungen und Beschreibungen zur Evaluation. Daher wird hier ausgeführt, was im Rahmen dieses Berichts unter „Evaluation“ verstanden wird. Dieses Unterkapitel ist dazu in folgende Abschnitte gegliedert:

- 2.3.1. Evaluation und Projektevaluation.
- 2.3.2. Ergebnisevaluation.
- 2.3.3. Prozessevaluation.

Eine allgemeine Einführung in die Projektevaluationen findet sich bei Wagner (2004).

2.3.1. Evaluation und Projektevaluation

Eine Evaluation ist zunächst zu verstehen als jegliche Form der Bewertung einer Sache (z.B. Scriven, 1991). Diese sehr allgemeine Definition belässt jedoch den Gegenstandsbereich von Evaluationen weitgehend offen und macht auch keine konkreten Aussagen darüber, mit welchen Methoden evaluiert werden kann. Aus diesem Grund bleibt bei einer fehlenden Spezifikation des Begriffs „Evaluation“ unklar, was gemeint ist. Zur Einordnung eines gegebenen Evaluationsprojekts ist es daher unerlässlich, genauer einzugrenzen, welche Evaluation durchgeführt wurde.

Eine Form der Evaluation ist die Projektevaluation (Synonyme: Programmevaluation oder Trainingsevaluation), d.h. Evaluationsgegenstand ist ein zu bewertendes Projekt oder Training – beispielsweise eine Maßnahme zur Prävention von Vorurteilen gegenüber ethnischen Minderheiten. Allerdings sind verschiedene methodische Ansätze denkbar, nach denen eine Maßnahme im Rahmen einer Projektevaluation bewertet werden kann. Daher spezifiziert auch der Begriff „Projektevaluation“ noch nicht hinreichend, um welche Art der Evaluation es sich genau handelt.

Von den verschiedenen methodischen Ansätzen zur Projektevaluation sind die beiden Ansätze der Ergebnis- und der Prozessevaluation besonders relevant. Im Rahmen einer Ergebnisevaluation (Synonyme: Wirkevaluation oder Effektevaluation) wird bewertet, ob das Training zu dem gewünschten Ergebnis führt, d.h., ob durch das Training die aufgestellten Trainingsziele auch wirklich erreicht werden. Im Kontext einer Prozessevaluation wird überprüft, wie das Training aus der Perspektive der TrainingsteilnehmerInnen und / oder aus der Perspektive der TrainingsleiterInnen bewertet wird. Ergebnis- und Prozessevaluation sind nicht isoliert zu betrachten, sie sollten sich im Idealfall ergänzen, um ein umfassendes Bild über das zu evaluierende Projekt zu erhalten.

Da die Evaluation des PiT-Gewaltpräventionstrainings als Ergebnis- und Prozessevaluation angelegt war und durchgeführt wurde, werden im Folgenden noch einige allgemeine Bemerkungen zur Ergebnis- und Prozessevaluation vorgenommen.

2.3.2. Ergebnisevaluation

In diesem Abschnitt wird auf notwendige Voraussetzungen einer angemessenen Ergebnisevaluation eingegangen. Dabei werden die folgenden Aspekte berücksichtigt:

1. Formulierung von messbaren Trainingszielen.
2. Konstruktion eines geeigneten Messinstruments.
3. Verwendung eines geeigneten Untersuchungsdesigns.

2.3.2.1. Trainingsziele

Eine Prämisse der Ergebnisevaluation ist, dass vor Programmbeginn durch die TrainingsinitiatorInnen eine Definition der Trainingsziele erfolgt. Eine weitere wichtige Voraussetzung im Zusammenhang mit den Trainingszielen ist, dass diese möglichst konkret formuliert werden. Nur in diesem Fall ist die für eine Ergebnisevaluation notwendige Bestimmung des Ausmaßes der Zielerreichung möglich. Handelt es sich um ein komplexes Ziel, so ist dieses in konkrete Teilziele zu zerlegen.

Ein Ziel ist konkret definiert, wenn es sich auf die Veränderung der Ausprägung in einem klar umgrenzten und prinzipiell messbaren Merkmal bezieht, d.h. auf den Zuwachs der Ausprägung in einem erwünschtem Merkmal (z.B. Zivilcourage) oder auf die Reduktion der Ausprägung in einem unerwünschtem Merkmal (z.B. Aggressionsneigung).

2.3.2.2. Messinstrument

Eine weitere Prämisse der Ergebnisevaluation ist der Einsatz eines Messinstruments. Dabei werden meist Fragebögen verwendet. Ein Fragebogen sollte so konstruiert sein, dass er Indikatoren enthält, mit denen die Ausprägung der in den Zielen enthaltenen Merkmale (z.B. Vorurteile) gemessen werden kann. Ein Indikator setzt sich in der Regel aus mehreren Items (Aussagen) zusammen (z.B. „Die deutsche Politik kümmert sich zu viel um die hier lebenden Ausländer“), mit denen die Erfassung der Merkmalsausprägung erfolgt. Die Befragten müssen dazu zu jedem Item – meist durch Ankreuzen vorgegebener Antwortkategorien – Stellung nehmen.

In der Auswertung wird später mit den erhobenen Daten überprüft, ob sich aus dem Indikator eine empirisch begründete Skala bilden lässt, die das intendierte Merkmal mit hinreichender methodischer Qualität erfasst. Anhand dieser Skalenwerte können dann Angaben zur Trainingseffektivität gemacht werden.

2.3.2.3. Untersuchungsdesign

Als dritte Voraussetzung der Ergebnisevaluation ist die Wahl eines geeigneten Untersuchungsdesigns zur Datenerhebung zu nennen. Ziel ist es, ein Untersuchungsdesign zu verwenden, das unter den gegebenen praktischen Rahmenbedingungen bestmöglichst sicherstellt, dass sich sichere Aussagen zur Trainingseffektivität machen lassen. Theoretisch sind viele verschiedene Untersuchungsdesigns denkbar, diese sind jedoch unterschiedlich gut zur Ergebnisevaluation geeignet. Im Folgenden werden drei mögliche Untersuchungsdesigns kurz geschildert.

a) Nachtest-Design

Eine Möglichkeit der Datenerhebung besteht darin, lediglich nach Durchführung des Trainings eine Messung der in den Zielen enthaltenen Merkmale (z.B. Aggressionsneigung) vorzunehmen. Dadurch lässt sich zwar der Status Quo der Merkmalsausprägung in der Trainingsgruppe ermitteln, es lässt sich aber nicht bestimmen, welche Auswirkung das Training auf diese Ausprägung hat. Hier wird eine wichtige Mindestanforderung an ein Untersuchungsdesign deutlich: Es muss eine Referenzbasis zur Verfügung stellen, anhand derer sich beurteilen lässt, ob ein Trainingseffekt vorliegt, d.h., ob es durch das Training in der Trainingsgruppe zu einer erwünschten Veränderung kommt.

b) Vortest-Nachtest-Design

Eine in der Praxis häufig verwendete Form der Gewinnung von Vergleichswerten besteht darin, die Trainingsgruppe nicht nur nach der Durchführung, sondern auch vor Beginn des Trainings mit dem Messinstrument zu befragen. Durch den Vergleich der Merkmalsausprägung zum ersten Messzeitpunkt (t_1) mit der Merkmalsausprägung zum zweiten Messzeitpunkt (t_2) lässt sich so eine Aussage über den Trainingseffekt ableiten. Für den Erfolg der Maßnahme würde diesbezüglich sprechen, wenn beispielsweise die Aggressionsneigung nach dem Training geringer als vor dem Training ausfällt. Allerdings ist bei einem derartigem Design zu bedenken, dass zwischen den beiden Messzeitpunkten auch andere Faktoren wirksam sein können (z.B. merkmalsbezogene Medienberichterstattung, merkmalsbezogene gesetzgeberische Maßnahmen oder Reifungsprozesse bei den Befragten), die zu einer Veränderung der Ausprägung der in den Zielen enthaltenen Merkmale führen, von der Trainingsmaßnahme aber unabhängig sind. Daher lässt sich mit einem einfachen Vortest-Nachtest-Design zwar eine Veränderung in der Merkmalsausprägung ermitteln, es bleibt jedoch unklar, ob diese Veränderung auch wirklich auf das Training zurückzuführen ist, oder durch andere Faktoren (mit)bedingt ist.

c) Vortest-Nachtest-Design mit Kontrollgruppe

Eine Möglichkeit zur Erhöhung der Sicherheit, mit der man einen Veränderungseffekt der Trainingsdurchführung zuschreiben kann, besteht im zusätzlichen Einbezug einer Kontrollgruppe. Abbildung 1 veranschaulicht ein solches Vortest-Nachtest-Design mit Kontrollgruppe exemplarisch.

Abbildung 1

Bedingung	T1		T2
Trainingsgruppe	Messung der Zielindikatoren Beispiel: - Aggression - ...	Durchführung des Projekts	Messung der Zielindikatoren Beispiel: - Aggression - ...
Kontrollgruppe	Messung der Zielindikatoren Beispiel: - Aggression - ...	-----	Messung der Zielindikatoren Beispiel: - Aggression - ...

Die Kontrollgruppe sollte sich nach Möglichkeit von der Trainingsgruppe nur darin unterscheiden, dass sie nicht an der Trainingsmaßnahme teilnimmt. Im Idealfall sollte daher die Zuordnung einzelner Untersuchungsteilnehmer zur Trainings- bzw. Kontrollgruppe nach dem Zufall erfolgen. Dies ist allerdings in der Praxis häufig nicht umsetzbar. Man spricht dann von einem „nicht-äquivalentem Kontrollgruppendesign“. Unterschiede zwischen den nicht-äquivalenten Trainings- und Kontrollgruppen sind in diesem Fall so gut wie möglich durch statistische Verfahren zu kontrollieren.

Nach der inferenzstatistischen Auswertung (siehe später) einer Erhebung mit einem Vortest-Nachtest-Design mit Kontrollgruppe ist dann mit ziemlicher Sicherheit auf eine positive Trainingswirkung zu schließen, wenn eins der beiden folgenden Ergebnisse vorliegt:

- a) In der Trainingsgruppe ist vom ersten zum zweiten Messzeitpunkt eine erwünschte Merkmalsveränderung zu beobachten (z.B. Absinken der Aggressionsneigung), während in der Kontrollgruppe eine erheblich schwächere erwünschte Veränderung, keine Veränderung oder eine unerwünschte Veränderung (z.B. Anstieg der Aggressionsneigung) zu beobachten ist.
- b) In der Trainingsgruppe ist keine oder eine schwache unerwünschte Merkmalsveränderung zu beobachten, während in der Kontrollgruppe eine erheblich stärkere unerwünschte Veränderung zu beobachten ist. In diesem Fall wird durch das Training eine Art „Puffer“ etabliert, der vor einer unerwünschten Veränderung durch andere Einflüsse schützen kann.

Neben den hier vorgestellten Designs – von denen nach Möglichkeit dass zuletzt genannte herangezogen werden sollte – gibt es noch komplexere Designs, die teilweise weitergehende Aussagen erlauben, jedoch erheblich aufwendiger sind und in der Praxis daher nur sehr selten angewendet werden können.

2.3.3. Prozessevaluation

Im Rahmen einer Prozessevaluation wird in der Regel untersucht, wie die TrainingsteilnehmerInnen und / oder die TrainerInnen das Training hinsichtlich verschiedener Aspekte subjektiv bewerten. Sie werden dazu beispielsweise mit einem Fragebogen befragt. Die Befragung findet in der Regel nach dem Ende der Trainingsdurchführung statt und kann mit geschlossenen und / oder offenen Fragen erfolgen. Bei geschlossenen Fragen werden bestimmte Antwortkategorien vorgegeben (z.B.: 1 = stimme gar nicht zu; 2 = stimme nicht zu; 3 = stimme zu; 4 = stimme voll zu). Die Befragten haben dann anzukreuzen, welche Antwortkategorie für sie zutrifft. Gegenstand einer derartigen Befragung mit geschlossenen Fragen können beispielsweise die folgenden Aspekte sein:

- a) Subjektiv wahrgenommener Trainingseffekt und Transfer des Trainingseffekts auf den Alltag.
- b) Subjektive Wahrnehmung der Methodik des Trainings (z.B. hinsichtlich der Zeitplanung oder hinsichtlich der bereitgestellten Materialien).
- c) Subjektive Wahrnehmung der Rahmenbedingungen des Trainings (z.B. hinsichtlich der Zusammenarbeit oder hinsichtlich der Arbeitsatmosphäre).

Offene Fragen geben keine Antwortkategorien vor, die Person hat vielmehr die Möglichkeit, in eigenen Stichworten auf die Frage zu antworten. In der Regel wird dabei nach folgenden Aspekten gefragt:

- a) (Besonders) effektive Elemente des Trainings.
- b) Verbesserungsvorschläge.

Offene Fragen haben den Vorteil, dass man auch Informationen zu neuen Aspekten erhält, die bisher nicht bedacht wurden. Diesem Vorteil steht jedoch der Nachteil einer erheblich aufwendigeren und stets mit einem gewissen subjektiven Interpretationsspielraum behafteten inhaltsanalytischen Auswertung gegenüber.

In den beiden folgenden Kapiteln wird zunächst auf die Ergebnis- und sodann auf die Prozessevaluation der PiT-Trainingsmaßnahmen eingegangen.

3. *Ergebnisevaluation*

Dieses Kapitel beinhaltet folgende Unterkapitel zur Ergebnisevaluation der im Rahmen von PiT-Hessen durchgeführten Trainingsmaßnahmen:

- 3.1. Methodik.
- 3.2. Auswertungsergebnisse.
- 3.3. Fazit zur Ergebnisevaluation.

3.1. Methodik

Das Unterkapitel Methodik ist folgendermaßen gegliedert:

- 3.1.1. Messinstrument.
- 3.1.2. Untersuchungsdesign.

Diese Abschnitte haben eine weitere Binnendifferenzierung und werden im Folgenden dargestellt.

3.1.1. Messinstrument

3.1.1.1. Allgemeines zum Messinstrument

Um die im Rahmen von PiT-Hessen durchgeführten Trainingsmaßnahmen evaluieren zu können, wurde zunächst von der Arbeitsgruppe Sozialpsychologie in Absprache mit der PiT-Lenkungsgruppe ein Fragebogen entwickelt. Das Messinstrument wurde von der AG Sozialpsychologie so angelegt, dass es Indikatoren für die Ausprägung der in den Trainingszielen thematisierten Fähigkeiten bzw. Merkmale (z.B. Zivilcourage) enthält – die Indikatoren machen die angesprochenen Fähigkeiten bzw. Merkmale messbar. Durch wiederholte Vorgabe des Messinstruments kann somit die Frage beantwortet werden, ob es durch das Training zum anvisierten Zuwachs der Ausprägungen dieser Fähigkeiten bzw. Merkmalen kommt (beispielsweise zu einer Stärkung der Zivilcourage).

Da der Fokus des Trainings nach Angaben der TrainingsinitiatorInnen auf die Verhaltensebene gerichtet ist, läge es zunächst nahe, alternativ oder ergänzend zur schriftlichen Befragung eine systematische Verhaltensbeobachtung und – beurteilung durchzuführen. Dieser Ansatz wurde anfänglich verfolgt, dann aber aufgrund von Umsetzungsproblemen fallen gelassen. Stattdessen wurde der Fragebogen in weiten Teilen möglichst verhaltensnah konzipiert. Das hier verwendete Messinstrument ist daher so konstruiert, dass die Ausprägung der in den Trainingszielen erwähnten Fähigkeiten bzw. Merkmale durch Indikatoren erfasst wird, die jeweils auf einer der folgenden vier Ebenen angesiedelt sind:

1. Ebene der Verhaltensintention.
2. Ebene der Situationsbewertung.
3. Einstellungsebene.
4. Ebene retrospektiver Verhaltensangaben.

Insgesamt wurden für das Messinstrument zwölf verschiedene Indikatoren entwickelt. Tabelle 3 gibt einen Überblick über die Zuordnungen zwischen Indikatorebenen, Indikatoren und Trainingszielen.

Tabelle 3

Indikatorebene	Indikator	Trainingsziel
1. Verhaltensintention.	1. Konflikteinstieg.	1. Entwicklung persönlicher Handlungsoptionen.
	2. Konfliktausstieg.	
	3. Konformität.	
	4. Zivilcourage.	
2. Situationsbewertung.		2. Hinsehen statt Wegesehen.
		3. Stärkung von Zivilcourage.
		4. Erlernen von „richtigem“ Helferverhalten in gewaltbesetzten Situationen.
		5. Frühzeitiges Einbinden von „Helfern“ in Problemsituationen.
	5. Gefahr.	6. Sensibilisierung für das frühzeitige Erkennen konflikthafter bzw. gefahrenträchtiger Situationen.
3. Einstellung.	6. Selbstwirksamkeit.	7. Weg von der passiven Opferrolle hin zum reflektierten Umgang mit gewaltbesetzten Situationen.
	7. Suche nach Lehrerunterstützung.	8. Vertrauensbildung/-Intensivierung gegenüber LehrerInnen in deren Rolle als Hilfsperson.
	8. Angst.	9. Reflexion der eigenen Ängste und Befürchtungen.
	9. Gewaltakzeptanz.	-
	10. Gewaltbereitschaft.	-
	11. Ausgeübte Aggression.	-
4. retrospektive Verhaltensangabe.	12. Erfahrene Angriffe.	-

Acht der zwölf Indikatoren (beispielsweise Konfliktausstieg oder Zivilcourage) wurden explizit zur Erfassung der in den Zielen aufgeführten Fähigkeiten bzw. Merkmale konzipiert. Die übrigen vier Indikatoren (Gewaltakzeptanz, Gewaltbereitschaft, ausgeübte Aggression und erfahrene Angriffe) sollen Einstellungen bzw. Verhaltensangaben erfassen, die mit den Trainingszielen in enger Verbindung stehen.

Jedem Indikator sind jeweils mehrere Items (Aussagen) zugeordnet, von denen angenommen wird, dass sie die Fähigkeit bzw. das entsprechende Merkmal messen können, das durch den Indikator gemessen werden soll.

In der Befragung werden den SchülerInnen die Items einzeln und nach Indikatoren geordnet vorgelegt, sie sollen durch Ankreuzen vorgegebener Antwortkategorien zu jedem Item Stellung nehmen. Abbildung 2 verdeutlicht das Prinzip anhand eines Ausschnitts aus dem Fragebogen:

Abbildung 2

Was glaubst du, wie du allgemein mit bedrohlichen Situationen oder Streit umgehen kannst?			
	nein	weiß nicht	ja
Ich komme auch mit überraschenden Streitsituationen klar.	1	2	3

In der Auswertung wird später für jede Schülerin bzw. für jeden Schüler pro Indikator ein Mittelwert der Angaben zu allen Items des Indikators berechnet. Dieser Mittelwert stellt die Ausprägung der Person in der durch den Indikator gemessenen Fähigkeit dar. Zur Interpretation der Mittelwerte gilt dann: Je höher der Wert, desto stärker ist die durch den Indikator zu messende Fähigkeit bei der Schülerin bzw. bei dem Schüler ausgeprägt.

An dieser Stelle sei auch erwähnt, dass die Antwortkategorien in der Befragung nicht in der Reihenfolge des oben genannten Beispiels präsentiert wurden, sondern umgekehrt (d.h. z.B. 1 = ja). Nach der Erhebung wurden sie in die hier und im übrigen Bericht verwendete Reihenfolge umkodiert, d.h. SchülerInnen, die in der Befragung mit „1“ geantwortet haben, wurde eine „3“ zugewiesen und SchülerInnen, die in der Befragung mit „3“ geantwortet haben, wurde eine „1“ zugewiesen. Die Werte lassen sich durch die Umkodierung besser interpretieren: Je höher der Wert, desto erwünschter die Ausprägung im Sinne der Trainingsziele.

Am Ende des Fragebogens finden sich Fragen zu soziodemographischen Merkmalen (Alter, Name der Schule, Klassenstufe, Schultyp, Geschlecht, Nationalität, Geburtsland, Geburtsland des Vaters, Geburtsland der Mutter, Muttersprache) sowie ein persönlicher Code, den die SchülerInnen zum ersten und zum zweiten Messzeitpunkt angeben mussten und der – bei adäquater Angabe – unter Gewährleistung der Anonymität für SchülerInnen, die den Fragebogen zu beiden Messzeitpunkten ausgefüllt haben, die Zusammenfügung dieser Fragebögen ermöglicht. Die gelungene Zusammenfügung der Fragebögen einer Schülerin oder eines Schülers ist Voraussetzung für die Aufnahme der SchülerInnen oder des Schülers in die Analyse der Trainingseffektivität (siehe später).

Auf den nächsten Seiten werden die einzelnen Indikatoren getrennt nach Indikator-ebene jeweils kurz vorgestellt. Diese Indikatoren wurden, fall nicht anders erwähnt, von der AG Sozialpsychologie eigens für die Evaluation von PiT-Hessen entwickelt.

3.1.1.2. Vorstellung der einzelnen Indikatoren

1. Ebene der Verhaltensintention

1. Einführende Bemerkungen

Ein Hauptziel für die Erstellung des Fragebogens war, der Erfassung von manifesten Verhaltensweisen so nah wie möglich zu kommen. Um Indikatoren auf der Ebene der Verhaltensintention in konkreten Konfliktsituationen erfassen zu können, wurde eine Szenario-Technik verwendet. Bei dieser Technik wird den SchülerInnen zunächst eine konkrete Situation zusammen mit der Aufforderung geschildert, sich in die diese Situation hinein zu versetzen:

- z.B. Du beobachtest im Bus, wie mehrere Typen, etwas älter als du, einen Jungen, so alt wie du, beschimpfen, bespucken und rumschubsen. Wie verhältst du dich in dieser Situation? (Szenario 2 des Indikators Zivilcourage).

Anschließend werden mehrere Items (Aussagen) aufgeführt, die jeweils eine mögliche Verhaltenreaktion auf diese Situation repräsentieren (z.B. Ich rufe den umstehenden Leuten zu: „Guckt nicht zu, helft lieber!“). Die SchülerInnen haben zu jeder Verhaltensweise anhand von drei vorgegebenen Antwortkategorien durch Ankreuzen anzugeben, ob das Verhalten auf sie zutrifft (1 = nein; 2 = vielleicht; 3 = ja).

Im Folgenden werden die vier auf der Ebene der Verhaltensintention verorteten Indikatoren näher beschrieben.

2. Konflikteinstieg (Indikator 1)

Dieser Indikator misst die Fähigkeit zur Vermeidung des Einstiegs in einen Konflikt. Er wird daher zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Entwicklung persönlicher Handlungsoptionen“ herangezogen.

Den SchülerInnen wird zunächst ein Szenario mit Konfliktgefahr beschrieben. Anschließend werden fünf Items dargeboten, die sich auf den Umgang mit der Konfliktgefahr beziehen (z.B. „Ich gehe mit großem Abstand an den Typen vorbei, ohne sie zu beachten“). Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter ist die Fähigkeit zur Vermeidung des Einstiegs in einen Konflikt.

3. Konfliktausstieg (Indikator 2)

„Konfliktausstieg“ misst die Fähigkeit zum Ausstieg aus einem (schon vorhandenem) Konflikt. Dieser Indikator wird ebenfalls zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Entwicklung persönlicher Handlungsoptionen“ verwendet.

Den Befragten wird eine Situation geschildert, in der ein Konflikt vorliegt. Im Anschluss an die Situationsschilderung werden zehn Items vorgegeben (z.B. „Ich stehe auf und gehe weg“). Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter ist die Fähigkeit zum Ausstieg aus einem (bereits vorhandenem) Konflikt.

4. Konformität (Indikator 3)

Der Indikator „Konformität“ misst das Ausmaß der Fähigkeit, sich der Dynamik von Gruppendruck zu entziehen. Laut Trainingsreader handelt es sich dabei um eine wesentliche Fähigkeit zum Umgang mit Konflikten. Aus diesem Grund wird auch „Konformität“ zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Entwicklung persönlicher Handlungsoptionen“ herangezogen.

Es werden vier Szenarien nacheinander geschildert, in denen die Gefahr besteht, sich Gruppendruck (entgegen der eigenen Ansichten und Bedürfnisse) zu beugen. Unterhalb der Szenarien werden Items aufgeführt, die spezifische Verhaltensweisen in der jeweiligen Situation beinhalten. Insgesamt sind diesem Indikator 16 Items zugeordnet (z.B. „Ich sage, dass ich nicht damit einverstanden bin, dass Daniel die Felswand hochklettern soll“). Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter ist die Fähigkeit, sich der Dynamik von Gruppendruck zu entziehen.

5. Zivilcourage (Indikator 4)

Der Begriff „Zivilcourage“ umschreibt ein Verhalten, bei dem die zivilcouragierte Person einer anderen Person oder Gruppe, die Opfer eines physischen oder psychischen Angriffs durch einen Aggressor wird, Hilfe leistet (Labuhn, Wagner, van Dick und Christ, 2004). Der Indikator „Zivilcourage“ misst das Ausmaß an Zivilcourage. Er wird zur Analyse des Trainingserfolgs hinsichtlich der vier Ziele „Hinsehen statt Wegsehen“, „Stärkung von Zivilcourage“, „Erlernen von „richtigem“ Helferverhalten in gewaltbesetzten Situationen“ und „frühzeitiges Einbinden von „Helfern“ in Problemsituationen“ verwendet.

Den SchülerInnen werden zwei Szenarien geschildert, in denen eine andere Person Opfer eines physischen oder psychischen Angriffs durch einen Aggressor wird. Im Anschluss an die beiden Situationen sind jeweils neun Items zu beantworten (z.B. „Ich tue so, als ob ich die Mädchen kennen würde, verwickle sie in ein Gespräch und gehe mit ihnen weg“). Die Items sind für beide Situationen identisch, sie wurden von Labuhn et al. (2004) entwickelt, teilweise wurde eine leichte thematische Anpassung der Fragen an die Trainingsinhalte vorgenommen. Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto stärker ist die Ausprägung des zivilcouragierten Verhaltens.

2. Ebene der Situationsbewertung

1. Einführende Bemerkungen

Auf dieser Ebene ist der Indikator „Gefahr“ verortet. Dieser Indikator bezieht sich nicht auf angemessene Verhaltensintentionen, sondern durch ihn soll untersucht werden, ob die SchülerInnen Situationen hinsichtlich ihres Gefahrenpotentials angemessen bewerten können.

2. Gefahr (Indikator 5)

„Gefahr“ soll das Ausmaß messen, in dem die SchülerInnen konfliktbehaftete bzw. gefahrenträchtige Situationen adäquat als gefährlich bewerten. Aus diesem Grund wird dieser Indikator zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Sensibilisierung für das frühzeitige Erkennen konflikthafter bzw. gefahrenträchtiger Situationen“ herangezogen. Es werden insgesamt vier Situationen geschildert.

- z.B. „Im Kaufhaus steht auf der Rolltreppe plötzlich ein unbekannter Typ neben dir, er ist ungefähr in deinem Alter und quatscht dich an. Erst fragt er dich nach der Uhrzeit und dann, ob er mal mit deinem Handy einen Freund anrufen könne, wegen einer Party, sein Handy sei gerade leer. Du könntest auch mitkommen auf die Party. Dir kommt der Typ seltsam und ziemlich aufdringlich vor“.

Die Situationen sind von den SchülerInnen jeweils hinsichtlich ihrer Gefährlichkeit anhand einer dreistufigen Skala (1 = nicht gefährlich; 2= weiß nicht; 3 = gefährlich) zu bewerten. Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter ist die Fähigkeit, konfliktbehaftete bzw. gefahrenträchtige Situationen adäquat als gefährlich zu bewerten.

3. **Einstellungsebene**

1. Einführende Bemerkungen

Die auf dieser Ebene verorteten Indikatoren beziehen sich auf Einstellungen. Einstellungen sind globale, latente Prädispositionen auf einen Reiz (z.B. soziale Probleme, Personen oder sonstige Einstellungsobjekte) in bestimmter Weise zu reagieren. Im Folgenden werden die fünf auf dieser Ebene angesiedelten Indikatoren kurz vorgestellt.

2. Selbstwirksamkeit (Indikator 6)

Der Begriff „Selbstwirksamkeit“ (engl. self-efficacy) beschreibt die individuellen Annahmen darüber, wie gut (oder schlecht) man in einer bestimmten Situation mit bestimmten Fähigkeiten, die man besitzt, und unter den gegebenen Umständen agieren kann (Bandura, 1997).

Der für dieses Messinstrument konstruierte Indikator „Selbstwirksamkeit“ ist so gestaltet, dass er das Ausmaß erfasst, in dem die befragten SchülerInnen von sich glauben mit bedrohlichen Situationen oder Streit umgehen zu können. „Selbstwirksamkeit“ wird daher zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Weg von der passiven Opferrolle hin zum reflektierten Umgang mit gewaltbesetzten Situationen“ herangezogen.

Den SchülerInnen werden vier Items vorgelegt (z.B. „In bedrohlichen Situationen weiß ich immer, wie ich mich verhalten soll“). Bei diesen Fragen wurde auf Items aus der Skala Selbstwirksamkeit von Schwarzer (1993) zurückgegriffen und diese an die Thematik des PiT-Programms angepasst.

Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter die Selbstwirksamkeit.

3. Suche nach Lehrerunterstützung (Indikator 7)

„Suche nach Lehrerunterstützung“ misst, inwiefern die Befragten in ernsthafteren Streitsituationen auf die Unterstützung von LehrerInnen zurückgreifen. Dieser Indikator wird zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Vertrauensbildung gegenüber LehrerInnen in deren Rolle als Hilfsperson“ verwendet.

Die SchülerInnen haben Antworten zu drei Items (z.B. „Bei ernsthaftem Streit in der Schule glaube ich, dass eine Lehrerin oder ein Lehrer helfen kann, den Streit zu schlichten“) abzugeben. Das Antwortformat ist dreistufig (1 = nein; 2 = vielleicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto ausgeprägter die Suche nach Lehrerunterstützung.

4. Angst (Indikator 8)

Der Indikator „Angst“ misst, in welchem Ausmaß die Befragten Ängste in Bedrohungs- und Streitsituationen als sinnvoll ansehen. „Angst“ wird zur Analyse des Trainingserfolgs hinsichtlich des Ziels „Reflexion der eigenen Ängste und Befürchtungen“ herangezogen.

Den Befragten werden vier wertende Aussagen zu Angst in Bedrohungs- oder Streitsituationen vorgegeben (z.B. „Ich glaube, Angst kann mir helfen zu erkennen, wie gefährlich eine Situation ist“). Das Antwortformat ist dreistufig (1 = nein; 2 = weiß nicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto sinnvoller wird Angst in Bedrohungs- und Streitsituationen angesehen.

Zusätzlich zu den bisher geschilderten Indikatoren wurden auf der Einstellungsebene noch zwei weitere Indikatoren herangezogen. Diese Indikatoren sind zwar nicht direkt mit den Trainingszielen verknüpft, sie stehen jedoch thematisch mit dem Training in enger Verbindung.

5. Gewaltakzeptanz (Indikator 9)

Der Indikator „Gewaltakzeptanz“ misst das Ausmaß der generellen Gewaltakzeptanz. Die SchülerInnen haben drei Items zu beantworten (z.B. „Ich finde es schlimm, wenn Leute sich prügeln“). Die Items gehen auf Sturzbecher (1997) zurück und wurden in ihren Formulierungen lediglich leicht abgeändert. Das Antwortformat ist dreistufig (1 = nein; 2 = weiß nicht; 3 = ja). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto geringer die Gewaltakzeptanz (d.h. desto erwünschter die Merkmalsausprägung im Sinne des Trainings).

6. Gewaltbereitschaft (Indikator 10)

„Gewaltbereitschaft“ misst das Ausmaß der Bereitschaft, selbst Gewalt anzuwenden. Den SchülerInnen werden vier Items vorgegeben (z.B. „Manchmal muss ich Gewalt einsetzen, um nicht den Kürzeren zu ziehen“). Die Items gehen auf Heitmeyer (1995) zurück und mussten in ihren Formulierungen nur geringfügig modifiziert werden. Das Antwortformat ist dreistufig (1 = ja; 2 = weiß nicht; 3 = nein). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto geringer die Gewaltbereitschaft (d.h. desto erwünschter die Merkmalsausprägung im Sinne des Trainings).

4. *Ebene retrospektiver Verhaltensangaben*

1. Einführende Bemerkungen

Auf dieser Ebene werden die SchülerInnen nach selbst ausgeführten und erlebten aggressiven Verhaltensweisen befragt. Die Erfassung dieser beiden Indikatoren ist ebenfalls wichtig, da sie mit den Trainingszielen in einem engen inhaltlichen Zusammenhang stehen.

2. Ausgeübte Aggression (Indikator 11)

Der Indikator „ausgeübte Aggression“ misst die Häufigkeit, mit der die Befragten (laut ihrer eigenen Angabe) in den vergangenen zwei Wochen konkrete aggressive Verhaltensweisen ausgeführt haben.

Die Schülerinnen werden nach der Häufigkeit des Ausführens von vier aggressiven Verhaltensweisen in den vergangenen zwei Wochen befragt (z.B. „Wie oft hast du in den letzten zwei Wochen andere geschlagen oder getreten?“). Das Antwortformat ist vierstufig und spiegelt Häufigkeitskategorien wieder (1 = mehr als 5 mal; 2 = 3-5 mal; 3 = 1-2 mal; 4 = kein mal). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto seltener wurden nach Angabe der Schülerin bzw. des Schülers von ihr bzw. ihm die angesprochenen aggressiven Verhaltensweisen ausgeführt.

3. Erfahrene Angriffe (Indikator 12)

„Erfahrene Angriffe“ misst die Häufigkeit, mit der in den vergangenen zwei Wochen bestimmtes aggressives Verhalten am eigenen Körper erfahren wurde.

Die SchülerInnen werden nach der Häufigkeit des Erfahrens von zwei Formen aggressiven Verhaltens befragt (z.B. „Wie oft bist du in den letzten zwei Wochen von anderen mit Worten angegriffen worden?“). Das Antwortformat ist vierstufig und spiegelt Häufigkeitskategorien wieder (1 = mehr als 5 mal; 2 = 3-5 mal; 3 = 1-2 mal; 4 = kein mal). Die Werte sind folgendermaßen zu interpretieren: Je höher der Wert, desto seltener wurden aggressive Angriffe erfahren.

Im nächsten Abschnitt wird näher geschildert, auf welche Art und Weise mit diesem Fragebogen Daten erhoben wurden, um eine Analyse des Trainingseffekts zu ermöglichen.

3.1.2. Untersuchungsdesign

Im Rahmen der allgemeinen Bemerkungen zur Ergebnisevaluation wurden drei verschiedene Untersuchungsdesigns vorgestellt. Dabei wurde auf die Nachteile eines Nachtest-Designs und eines einfachen Vortest-Nachtest-Designs eingegangen. Im Rahmen der Ergebnisevaluation von PiT-Hessen wurden die Daten nach dem Prinzip eines Vortest-Nachtest-Designs mit nicht-äquivalenter Kontrollgruppe erhoben. Wie bereits weiter oben erwähnt wurde, beteiligten sich die 16 Schulen des Modellprojekts Frankfurt-Offenbach mit jeweils einer bis fünf Trainings- und einer bis zwei Kontrollklassen an der Untersuchung. Dabei handelt es sich um sechs Gesamtschulen, vier Gymnasien, zwei Hauptschulen, zwei Realschulen und zwei Haupt- und Realschulen. In der Mehrzahl der Projektschulen entsprach eine Trainingsklasse jeweils einer Klassenstufe, in zwei Schulen stellte die Trainingsgruppe eine speziell konstruierte Arbeitsgruppe dar. Die Trainingsklassen entstammten den Jahrgangsstufen fünf und sieben bis neun, die Kontrollklassen den Jahrgängen sechs bis neun. Die Auswahl der Kontrollklassen erfolgte so, dass nach Möglichkeit Klassen aus derselben Jahrgangsstufe wie die Trainingsklassen gewählt wurden. War dies nicht möglich, weil z.B. der gesamte Jahrgang trainiert wurde, so sollte jeweils eine Klasse aus einem älteren und einem jüngeren Jahrgang ausgewählt werden.

Die Trainingsklassen nahmen am Training teil, sie werden im Folgenden zusammenfassend als „Trainingsgruppe“ bezeichnet. Die Trainingsmaßnahme umfasste laut Trainings-Reader vier Bausteine (Was ist Gewalt, Konflikteinstieg, Konfliktausstieg, Helferverhalten). Zu jedem Baustein wurden mehrere Übungen durchgeführt. Sowohl die Auswahl der Übungen, als auch der zeitliche Rahmen (Anzahl der Trainingsstunden insgesamt) und die Festlegung des Trainingssettings (z. B. einzelne Schulstunden vs. einmalige Projektwoche) erfolgte in den einzelnen Schulen individuell verschieden, so dass die Trainings hinsichtlich dieser Aspekte sehr heterogen waren. Allen Trainings war jedoch gemeinsam, dass sämtliche vier Bausteine behandelt wurden und die leitenden Teams sich in ihrer Zusammensetzung glichen: In jeder Schule gab es ein PiT-Team, welches das Training gemeinsam durchführte und sich aus jeweils mindestens zwei Mitgliedern des Lehrkörpers, der Polizei und der Jugendhilfe zusammensetzte.

Die Kontrollklassen nahmen nicht am Training teil, ihre Werte wurden lediglich als Referenz mit erhoben. Diese Klassen werden im Folgenden zusammenfassend als „Kontrollgruppe“ bezeichnet.

Das im vorhergehenden Abschnitt vorgestellte Messinstrument wurde sowohl der Trainingsgruppe als auch der Kontrollgruppe zweimal vorgegeben. Die Erhebung zum ersten Messzeitpunkt (t1) erfolgte vor der Trainingsdurchführung zu Beginn des Schuljahrs 2004/2005, die Erhebung zum zweiten Messzeitpunkt (t2) erfolgte gegen Ende des Schuljahrs 2004/2005 nach Abschluss der Trainingsdurchführung. Abbildung 4 gibt einen schematischen Überblick über das Untersuchungsdesign.

Abbildung 3

Bedingung	T1		T2
Trainingsgruppe	Messung der Zielindikatoren - Konf.einstieg - Zivilcourage - ...	Durchführung von PIT-Hessen	Messung der Zielindikatoren - Konf.einstieg - Zivilcourage - ...
Kontrollgruppe	Messung der Zielindikatoren - Konf.einstieg - Zivilcourage - ...	-----	Messung der Zielindikatoren - Konf.einstieg - Zivilcourage - ...

Im nächsten Unterkapitel werden die Ergebnisse der mit den erhobenen Daten vorgenommenen Auswertungen geschildert. Dabei werden auch die Auswertungsmethoden erläutert.

3.2. Auswertungsergebnisse

Dieses Unterkapitel ist in folgende Abschnitte gegliedert:

- 3.2.1. Angaben zur Analysestichprobe.
- 3.2.2. Empirische Skalenkonstruktion – Struktur und Qualität der Indikatoren.
- 3.2.3. Effektivität des Trainings – Ergebnisse der inferenzstatistischen Verfahren.

Diese Abschnitte haben eine weitere Binnendifferenzierung und werden im Folgenden dargestellt.

3.2.1. Angaben zur Analysestichprobe

3.2.1.1. Voraussetzungen zur Aufnahme in die Analysestichprobe

In der statistischen Auswertung zur Trainingseffektivität werden intraindividuelle Veränderungswerte von der Erstmessung (vor der Trainingsdurchführung) zur Zweitmessung (nach der Trainingsdurchführung) analysiert (siehe später). Aus diesem Grund können nur SchülerInnen in die Analyse aufgenommen werden, für die zu beiden Messzeitpunkten Daten vorliegen, und für die diese Daten anhand des Codes verknüpfbar sind. Es ergeben sich also zwei Bedingungen für die Aufnahme einer Schülerin bzw. eines Schülers in die Analysestichprobe:

1. Der Fragebogen wurde zu beiden Messzeitpunkten ausgefüllt.
2. Die beiden Fragebögen können miteinander verknüpft werden.

3.2.1.2. Kennzeichnung der Analytestichprobe

Das Modellprojekt wurde – wie bereits mehrfach erwähnt – an 16 Schulen durchgeführt. Zwei dieser 16 Schulen konnten nicht in die statistischen Analysen zur Ermittlung der Trainingseffektivität einbezogen werden, da der AG Sozialpsychologie aus diesen Schulen nur zu einem Messzeitpunkt ausgefüllte Fragebogen vorliegen.

Tabelle 4 gibt Aufschluss über die am Evaluationsprojekt beteiligten SchülerInnen. Die Häufigkeiten werden getrennt nach Trainings- und Kontrollgruppe dargestellt.

Tabelle 4

Bedingung	SchülerInnen zu Messzeitpunkt 1	SchülerInnen zu Messzeitpunkt 2	Analysestichprobe
Trainingsgruppe	817	743	551
Kontrollgruppe	543	556	372
Gesamt	1360	1299	923

In der zweiten Spalte werden die Anzahlen der SchülerInnen aufgeführt, für die zum 1. Messzeitpunkt ein ausgefüllter Fragebogen vorliegt. Dabei handelt es sich um 817 SchülerInnen in der Trainingsgruppe und 543 SchülerInnen in der Kontrollgruppe. Die dritte Spalte gibt die Häufigkeiten derjenigen SchülerInnen an, für die zum zweiten Messzeitpunkt ein ausgefüllter Fragebogen vorliegt. Dies sind 743 SchülerInnen in der Trainingsgruppe und 556 SchülerInnen in der Kontrollgruppe. Dabei kann es sich um SchülerInnen handeln, für die auch zum ersten Messzeitpunkt ein ausgefüllter Fragebogen vorliegt, dies muss aber nicht der Fall sein. In der vierten Spalte sind schließlich die Häufigkeitsangaben zur Analysestichprobe aufgeführt. Auf diese Stichprobe wird bei allen weiteren Analysen zurückgegriffen. Die Verknüpfung der beiden Fragebogen erfolgte anhand des Codes, der am Ende des Fragebogens anzugeben war. Die Analysestichprobe umfasst 551 SchülerInnen in der Trainingsgruppe und 372 SchülerInnen in der Kontrollgruppe.

An dieser Stelle sei auch erwähnt, dass die oben genannten Fallzahlen der Analysestichprobe nicht den späteren Fallzahlen in der eigentlichen Datenauswertung entsprechen. Die Analysestichprobe wurde dort nochmals etwas reduziert (Fallzahlen siehe später), da das statistische Verfahren SchülerInnen ausschloss, die in mindestens einem Indikator nicht die Mehrheit der Items beantwortet haben.

Die Fallzahlen der Analysestichprobe sind bereits hier spürbar geringer als die Fallzahlen in Spalte 2 und 3. Dieser Schwund lässt sich durch folgende Aspekte erklären. Zum einen haben mehrere TeilnehmerInnen nicht zu beiden Messzeitpunkten einen Fragebogen ausgefüllt. Dies kann auf fehlende Anwesenheit oder auf Ausfüllungsverweigerung zu einem der beiden Messzeitpunkte zurückzuführen sein. Diese SchülerInnen werden daher nur unter Spalte 2 oder unter Spalte 3 der Tabelle aufgeführt. Ein weiterer Grund ist in der Tatsache zu sehen, dass mehrere TeilnehmerInnen den Verknüpfungscodes nicht oder nicht einheitlich ausgefüllt haben. Die Fragebögen dieser SchülerInnen konnten daher nicht einander zugeordnet und somit nicht weiter berücksichtigt werden.

Im nächsten Unterpunkt wird eine kurze Beschreibung der Analysestichprobe hinsichtlich einiger soziodemografischer Variablen vorgenommen.

3.2.1.3. Soziodemografische Beschreibung der Analysestichprobe

An dieser Stelle soll die Analysestichprobe getrennt nach Trainings- und Kontrollgruppe in den soziodemografischen Variablen Alter, Geschlecht und Migrationshintergrund beschrieben werden. Die Variable Migrationshintergrund ist dichotom: Migrationshintergrund gegeben vs. Migrationshintergrund nicht gegeben. Ein Migrationshintergrund war gegeben, wenn entweder der Schüler bzw. die Schülerin oder mindestens ein Elternteil nicht in Deutschland geboren wurde.

Tabelle 5 zeigt die Kennwerte von Trainings- und Kontrollgruppe in den berücksichtigten Variablen.

Tabelle 5

Variable	Kennwert	Trainingsgruppe	Kontrollgruppe
Alter	Mittelwert	12,62	12,69
	Standardabw.	2,19	1,40
	Anzahl (N)	548	372
Geschlecht	Anzahl weibl.	281	196
	Anzahl männl.	257	160
	Gesamt	538	356
Migrationshintergrund	Anzahl ja	329	209
	Anzahl nein	222	161
	Gesamt	551	370

Bei der deskriptiven Inspektion der Variablenwerte fallen zunächst keine erheblichen Differenzen zwischen den beiden Gruppen auf. Die Unterschiede zwischen Trainings- und Kontrollgruppe wurden dennoch auf Signifikanz getestet (zum Signifikanzbegriff siehe später), da (bereits geringe) systematische Differenzen in diesen Variablen den Schluss von eventuellen Unterschieden in der Ausprägung der Indikatoren auf einen Effekt des Trainings gefährden könnten.

Zur Signifikanztestung der Altersunterschiede wurde ein t-Test für unabhängige Stichproben berechnet. Die Differenz ist auf dem 5%-Niveau nicht signifikant ($t(914,81) = -0,60$; $p = 0,55$). Zur Prüfung der Unterschiede zwischen der Trainingsgruppe und der Kontrollgruppe in den Variablen Geschlecht und Migrationshintergrund wurde jeweils ein χ^2 -4-Felder-Test berechnet. Auch diese Signifikanztests ergeben, dass keine bedeutsamen Differenzen zwischen Trainingsgruppe und Kontrollgruppe vorliegen.

3.2.2. Empirische Skalenkonstruktion – Struktur und Qualität der Indikatoren bzw. Skalen

In diesem Unterkapitel werden die Ergebnisse zur Struktur und Qualität der Indikatoren bzw. Skalen berichtet. Dabei werden auch die herangezogenen statistischen Verfahren kurz erläutert. Das Unterkapitel ist im Einzelnen in die folgenden Abschnitte gegliedert:

- 3.2.2.1. Einführende Bemerkungen.
- 3.2.2.2. Ergebnis der Faktorenanalysen – Struktur der Indikatoren.
- 3.2.2.3. Ergebnis der Reliabilitätsanalysen – Qualität der Skalen.

3.2.2.1. Einführende Bemerkungen

1. Ausgangssituation

Im zweiten Kapitel wurden im Abschnitt zur Beschreibung des Messinstruments die verschiedenen Indikatoren des Messinstruments vorgestellt. Die Indikatoren sollten jeweils mit mehreren Items (Aussagen) eine Fähigkeit oder ein Merkmal messen. Bevor analysiert wird, ob es in den gemessenen Merkmalsausprägungen durch das Training zu Veränderungen kam, muss empirisch geprüft werden, ob die Items der Indikatoren, wie vermutet nur die jeweilige Fähigkeit bzw. das jeweilige Merkmal erfassen und ob aus den Indikatoren daher jeweils eine empirisch begründete Skala konstruiert werden kann. In diesem Rahmen wurden zunächst Faktorenanalysen (explorative Faktorenanalysen mit Extraktion nach der Hauptkomponentenmethode) und anschließend Reliabilitätsanalysen (im Sinne der internen Konsistenz) durchgeführt. Diese statistischen Verfahren dienen der Überprüfung der Struktur der Items eines Indikators (Faktorenanalysen) sowie der Berechnung der Reliabilität der – anhand der Ergebnisse der Faktorenanalysen – vorgeschlagenen Skalen (Reliabilitätsanalysen).

2. Zum Verfahren der Faktorenanalyse

Der Begriff „Faktorenanalyse“ bezeichnet eine Gruppe von komplexen statistischen Verfahren, die ausgehend von den Antworten der Personen auf die Items eines Indikators überprüfen, ob die Items ähnliche Informationen erfassen (einfaktorielle Lösung), und sie daher wie intendiert zu einer empirisch begründeten gemeinsamen Skala zusammengefasst werden dürfen, oder, ob die Items eines Indikators teilweise unterschiedliche Informationen erfassen (mehrfaktorielle Lösung), und der Indikator daher in mehrere separate Skalen zu zerlegen ist.

3. Zum Verfahren der Reliabilitätsanalyse

Die Reliabilität einer Skala ist ein Gütekriterium der klassischen Testtheorie. Dieser Begriff beschreibt allgemein die formale Genauigkeit, mit der eine Skala das intendierte Merkmal misst. Es gibt verschiedene Zugangsweisen zur empirischen Bestimmung der Reliabilität. Die üblichste Methode ist die Berechnung der internen Konsistenz der Skala. Die interne Konsistenz bezieht sich darauf, wie homogen die Items einer – zuvor durch die Faktorenanalyse vorgeschlagenen – Skala jeweils das intendierte Merkmal erfassen, mit anderen Worten, wie stark die unterschiedlichen Items einer Skala in den Augen der Befragten dasselbe Messen. Cronbach's Koeffizient α ist ein Index dieser Homogenität, er kann üblicherweise Werte in einem Spektrum von 0 bis 1 annehmen. Dabei gilt die Faustregel: Je höher α , desto homogener die Skala. Da Cronbach's α jedoch nicht nur von der Homogenität einer Skala, sondern auch von der Itemanzahl der Skala abhängt, lassen sich keine exakten Vorgaben machen, ab welchem Wert eine Skala als hinreichend homogen zu bezeichnen ist. Als ungefährender Richtwert gilt jedoch, dass Skalen für die α kleiner als 0,6 ist, nicht hinreichend reliabel sind. Die Aussagen aus diesen Skalen sind erheblich fehlerbehaftet, so dass sie aus den folgenden Analysen ausgeschlossen werden sollten. Das Verfahren liefert auch Angaben dazu, ob sich die interne Konsistenz einer Skala durch die Elimination von Einzelitems, die Informationen erfassen, die zu den anderen Items nicht hinreichend ähnlich sind, verbessern lässt.

3.2.2.2. Ergebnis der Faktorenanalysen – Struktur der Indikatoren

Zur Überprüfung der Struktur der einzelnen Indikatoren wurde mit den Items jedes der zwölf Indikatoren jeweils eine separate Faktorenanalyse (explorative Faktorenanalyse mit Extraktion nach der Hauptkomponentenmethode) berechnet. Für diejenigen Indikatoren, die mehrere Situationen implizieren (Konformität und Zivilcourage) wurde jeweils eine Analyse über alle Situationen berechnet. Die Verfahren wurden getrennt für die Trainingsgruppe und die Kontrollgruppe und innerhalb jeder Gruppe sowohl zum ersten Messzeitpunkt als auch zum zweiten Messzeitpunkt berechnet.

Die berechneten Faktorenanalysen ergaben für elf der zwölf Indikatoren in den unterschiedlichen Datensubstichproben eine einfaktorielle Lösung, d.h. die Items des Indikators messen das gleiche Merkmal, so dass jeweils für jeden dieser Indikatoren die Konstruktion nur einer Skala gerechtfertigt ist. Namen und Inhalte dieser Skalen entsprechen den Namen und Inhalten der ursprünglichen Indikatoren. Auf die Ergebnisse dieser Analysen wird im Folgenden nicht weiter eingegangen.

Die Faktorenanalysen mit den Items des Indikators Zivilcourage ergaben in jeder der vier Datensubstichproben jeweils, dass eine zweifaktorielle Lösung am geeignetsten ist, d.h. die Items dieses Indikators lassen sich zwei unterschiedlichen Skalen zuordnen. Diese Skalen wurden als Zivilcourage 1 („Kommunikation mit Umstehenden“) und Zivilcourage 2 („direktes Eingreifen“) bezeichnet. Zivilcourage1 bezieht sich auf die Kontaktaufnahme mit umstehenden Personen mit dem Ziel, diese zum Eingreifen zu motivieren. Diese Skala wurde daher „Kommunikation mit Umstehenden“ benannt.

Demgegenüber enthält Zivilcourage 2 Items, die sich auf direktes Eingreifen zur Hilfe des Opfers beziehen. Diese Skala trägt daher den Namen „direktes Eingreifen“. Tabelle 6 stellt dar, welche Items den beiden neuen Skalen durch die Faktorenanalyse zugeordnet wurden.

Tabelle 6

Skala	Items
1. Zivilcourage 1 („Kommunikation mit Umstehenden“)	<ol style="list-style-type: none"> 1. Ich versuche, andere Leute anzusprechen, um sie zum Eingreifen zu bringen. (Szenario 1) 2. Ich versuche, andere Leute anzusprechen, um sie zum Eingreifen zu bringen. (Szenario 1) 3. Ich rufe den umstehenden Leuten zu: „Guckt nicht zu, helft lieber!“ (Szenario 1) 4. Ich rufe den umstehenden Leuten zu: „Guckt nicht zu, helft lieber!“ (Szenario 2) 5. Ich fordere die Leute, die sich in meiner Nähe befinden auf, mit mir gemeinsam etwas gegen das Verhalten der Typen zu tun. (Szenario 1) 6. Ich fordere die Leute, die sich in meiner Nähe befinden auf, mit mir gemeinsam etwas gegen das Verhalten der Typen zu tun. (Szenario 2)
2. Zivilcourage 2 („direktes Eingreifen“)	<ol style="list-style-type: none"> 1. Ich tue so, als ob ich den Jungen kennen würde, verwickle ihn in ein Gespräch und gehe mit ihm weg. (Szenario 1) 2. Ich tue so, als ob ich den Jungen kennen würde, verwickle ihn in ein Gespräch und gehe mit ihm weg. (Szenario 2) 3. Ich spreche den Jungen an und frage, was los ist. (Szenario 1) 4. Ich spreche den Jungen an und frage, was los ist. (Szenario 2) 5. Ich ziehe den Jungen von den Typen weg. (Szenario 1) 6. Ich ziehe den Jungen von den Typen weg. (Szenario 2)

Die beiden Skalen bestehen jeweils aus sechs Items. Dabei wird jedes Item zweimal aufgeführt, da es von den Befragten sowohl hinsichtlich Szenario 1 als auch hinsichtlich Szenario 2 beantwortet wurde.

Zusammenfassend lässt sich an dieser Stelle festhalten, dass durch die Faktorenanalysen dreizehn Skalen vorgeschlagen werden. Auf den nächsten Seiten wird dargestellt, ob diese Skalen hinreichend reliabel sind, um sie in die Auswertung des Trainingseffekts miteinzubeziehen.

3.2.2.3. Reliabilitätsanalysen – Qualität der Skalen

Es wurden Reliabilitätsanalysen im Sinne der internen Konsistenz berechnet. Auf die Grundidee des Verfahrens wurde bereits eingegangen. Tabelle 7 enthält die internen Konsistenzen (Cronbach's Koeffizient α) der Skalen, getrennt für die Trainingsgruppe (TG) und die Kontrollgruppe (KG) zu den beiden Messzeitpunkten t1 (vor Durchführung des Trainings) und t2 (nach Durchführung des Trainings). Diese internen Konsistenzen wurden durch Ausschluss ungeeigneter Items, die Informationen erfassen, die zu den anderen Items der Skala nicht hinreichend ähnlich sind, optimiert.

Tabelle 7

Skala	Item- zahl	Cronbachs α			
		t1		t2	
		TG	KG	TG	KG
1. Konflikteinstieg	3	.59	.68	.60	.69
2. Konfliktausstieg	5	.60	.62	.69	.67
3. Konformität	9	.81	.80	.81	.82
4. Zivilcourage 1 - Kommunik. mit Umst.	6	.88	.87	.88	.90
5. Zivilcourage 2 - direktes Eingreifen	6	.83	.81	.83	.83
6. Gefahr	4	.49	.47	.57	.56
7. Selbstwirksamkeit	4	.62	.61	.63	.66
8. Suche nach Lehrerunterstützung	3	.80	.80	.78	.82
9. Angst	4	.66	.72	.72	.73
10. Gewaltakzeptanz	3	.43	.57	.54	.58
11. Gewaltbereitschaft	4	.73	.76	.77	.77
12. Ausgeübte Aggression	4	.60	.64	.73	.62
13. Erfahrene Angriffe	2	.64	.62	.66	.63

Aus Tabelle 7 wird ersichtlich, dass elf der dreizehn Skalen als hinreichend reliabel klassifiziert werden können, ihre Items erfassen demzufolge hinreichend homogen das anvisierte Merkmal. Die Reliabilitäten einer Skala unterscheiden sich zwischen den Messzeitpunkten und zwischen den beiden Gruppen nicht wesentlich.

Die Reliabilitäten der beiden Skalen Gefahr und Gewaltakzeptanz sind durchweg – zum Teil deutlich – geringer als 0,60, d.h. die beiden Skalen erfassen das jeweilige Merkmal nicht mit hinreichender Genauigkeit. Die Kennwerte sind durch die Elimination weiterer Items nicht zu verbessern. Die Skalen Gefahr und Gewaltakzeptanz müssen daher aus den folgenden Analysen zur Trainingseffektivität ausgeschlossen werden. An dieser Stelle sei jedoch erwähnt, dass explorativ dennoch einige Analysen mit den Items dieser Skalen berechnet wurden, diese führten jedoch zu keinen nennenswerten Ergebnissen.

3.2.2.4. Fazit zur empirischen Skalenkonstruktion

Aufgrund der Ergebnisse der Faktoren- und Reliabilitätsanalysen kann gefolgert werden, dass das Messinstrument sowohl in der Trainingsgruppe als auch in der Kontrollgruppe die Bildung von elf reliablen Skalen gestattet. Im Folgenden sind diese Skalen noch einmal aufgelistet:

1. Konflikteinstieg.
2. Konfliktausstieg.
3. Konformität.
4. Zivilcourage 1 - Kommunikation mit Umstehenden.
5. Zivilcourage 2 - direktes Eingreifen.
6. Angst.
7. Selbstwirksamkeit.
8. Suche nach Lehrerunterstützung.
9. Gewaltbereitschaft .
10. Ausgeübte Aggression.
11. Erfahrene Angriffe.

Für diese Skalen können zu den verschiedenen Messzeitpunkten jeweils getrennte Mittelwerte für die Trainingsgruppe und die Kontrollgruppe berechnet werden, anhand derer die Wirksamkeit des Trainings bestimmt werden kann. Dieser Aspekt wird im folgenden Abschnitt zusammen mit den Ergebnissen näher ausgeführt.

3.2.3. Effektivität des Trainings

Unter diesem Abschnitt werden die Ergebnisse der Auswertungen zur Trainingseffektivität vorgestellt. Die Darstellung ist nach folgenden Punkten gegliedert.

- 3.2.3.1. Einführende Bemerkungen.
- 3.2.3.2. Effektivität des Trainings I – Ergebnis der multivariaten Analyse.
- 3.2.3.3. Effektivität des Trainings II – Ergebnis der univariaten Analysen.
- 3.2.3.4. Effektivität des Trainings III – Einfluss weiterer Faktoren.

3.2.3.1. Einführende Bemerkungen

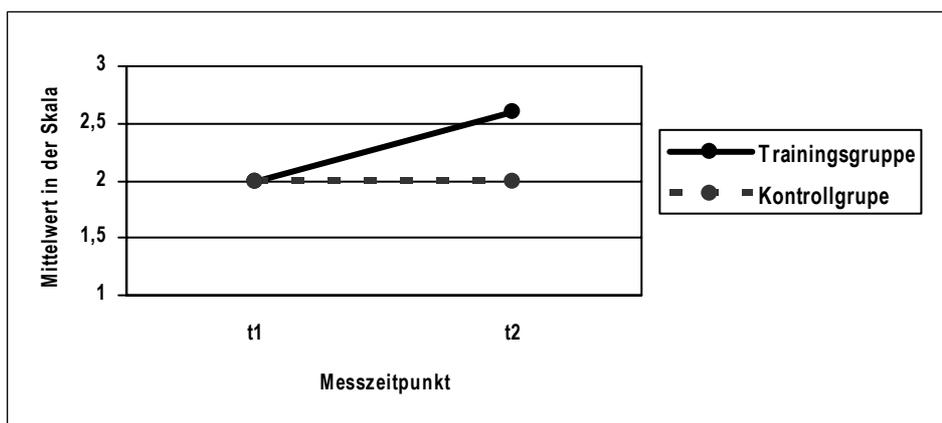
An dieser Stelle werden zunächst einige einführende Bemerkungen gemacht, sie beziehen sich auf drei Aspekte: Wechselwirkung, Konzept der Signifikanztestung sowie die Stärke signifikanter Effekte.

1. Bemerkungen zum Begriff der Wechselwirkung

Ein positiver Effekt des Trainings hinsichtlich eines der Trainingsziele würde sich in einer Wechselwirkung (Synonym: Interaktion) der Faktoren Bedingung (Trainingsgruppe vs. Kontrollgruppe) und Zeitpunkt der Messung (vor der Trainingsdurchführung vs. nach der Trainingsdurchführung) auf die Ausprägung in der Skala zeigen, die zur Messung des betreffenden Ziels herangezogen wird. Unter einer derartigen Wechselwirkung ist zu verstehen, dass sich die Merkmalsausprägung innerhalb der Trainingsgruppe vom ersten Messzeitpunkt zum zweiten Messzeitpunkt anders verändert als innerhalb der Kontrollgruppe. Dabei ist jedoch zu bedenken, dass nicht jede Wechselwirkung für einen positiven Trainingseffekt spricht. Es sind zwei Wechselwirkungskonstellationen denkbar, die einen Trainingseffekt indizieren:

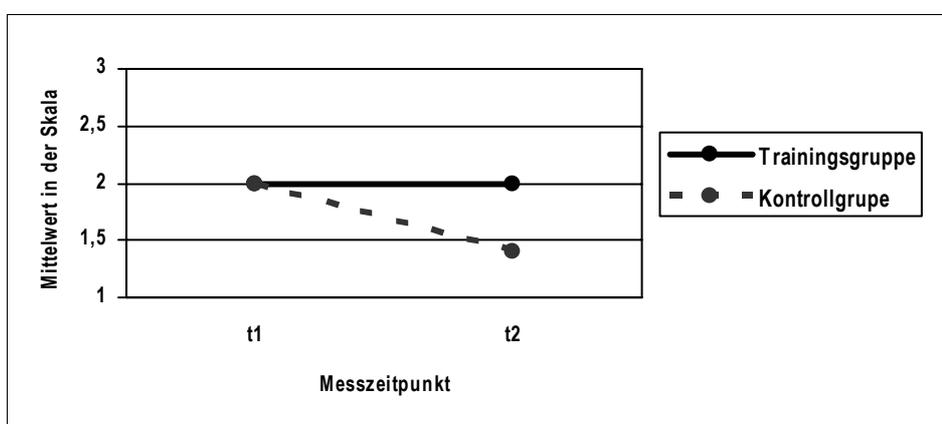
- a) die Merkmalsausprägung verändert sich in der Trainingsgruppe vom ersten zum zweiten Messzeitpunkt in der gewünschten Richtung (z.B. Zunahme von Zivilcourage), während es in der Kontrollgruppe über die Zeit hinweg zu einer deutlich geringeren Merkmalsveränderung in die gewünschte Richtung, zu keiner Merkmalsveränderung oder zu einer Merkmalsveränderung in die unerwünschte Richtung kommt. Abbildung 4 verdeutlicht dies exemplarisch.

Abbildung 4



- b) Die Merkmalsausprägung verändert sich in der Trainingsgruppe nicht oder nur geringfügig in die unerwünschte Richtung, während es in der Kontrollgruppe zu einer (erheblichen) Merkmalsveränderung in die unerwünschte Richtung kommt. In diesem Fall ist davon auszugehen, dass das Training einen indirekten Effekt hat, in dem es einen „Puffer“ etabliert, der gegen eine, durch andere Faktoren bedingte, Verschlechterung abfedert. Abbildung 5 verdeutlicht eine derartige Pufferwirkung exemplarisch.

Abbildung 5



2. Bemerkungen zum Konzept der Signifikanztestung

Im Rahmen einer Ergebnisevaluation möchte man eine Aussage darüber erhalten, wie wirksam das Training für die anvisierte Zielgruppe ist (beispielsweise für Jugendliche der siebten und achten Jahrgangsstufe). In der Praxis ist jedoch die Untersuchung der gesamten Zielgruppe (im Folgenden als Zielpopulation bezeichnet) in der Regel versperrt. Man muss das Training vielmehr anhand der Daten einer aus der Zielpopulation gezogenen Stichprobe evaluieren (beispielsweise Jugendliche der siebten und achten Jahrgangsstufe in einigen Projektschulen). Aus den Daten dieser Stichprobe soll dann eine Aussage für die – eigentlich interessierende – Zielpopulation abgeleitet werden. Dies kann nicht direkt 1:1 aus den Stichprobendaten erfolgen, da diese stets zufälligen Schwankungen unterliegen und nicht deckungsgleich mit den Daten der Zielpopulation sind. Würde man lediglich die Werte der Stichprobe betrachten und eine Aussage über die Trainingswirkung für die Zielpopulation machen, so wäre die Richtigkeit der Aussage hochgradig ungewiss. Aus diesem Grund sind die anhand einer Stichprobe ermittelten Daten stets mit einem inferenzstatistischen Verfahren auf Signifikanz zu testen. Es gibt verschiedene Arten solcher Verfahren (z.B. Verfahren aus der Gruppe der Varianzanalyse), die jeweils für bestimmte Fragestellungen und Datenkonstellationen geeignet sind. Sie beruhen alle auf dem Prinzip der Signifikanztestung. Ein Signifikanztest ist ein statistisches Verfahren, das die Wahrscheinlichkeit ermittelt, mit der die an der Stichprobe ermittelten Unterschiede oder noch extremere Unterschiede als die erhaltenen auftreten können, wenn das Training in der Zielpopulation keinen Effekt hat. Man nennt diese Wahrscheinlichkeit auch Irrtumswahrscheinlichkeit p , da sie der Wahrscheinlichkeit entspricht, mit der man sich irren würde, wenn man bei den gegebenen Unterschieden in der Stichprobe davon ausgehen würde, dass das Training in „Wahrheit“ (also für die Zielpopulation) einen Effekt hat. Als willkürliche Konvention hat sich eingebürgert, bei einer Irrtumswahrscheinlichkeit größer 0,05 davon auszugehen, dass ein Effekt nur zufälliger Natur und „in Wahrheit“ (also in der Zielpopulation) nicht vorhanden ist. Erst bei einer Irrtumswahrscheinlichkeit kleiner oder gleich 0,05 ist mit hinreichender Sicherheit davon auszugehen, dass ein Effekt „in Wahrheit“ vorhanden ist. Man spricht in diesem Fall von einem „signifikanten Ergebnis“. Das Ergebnis wird als hochsignifikant bezeichnet, wenn die Irrtumswahrscheinlichkeit kleiner oder gleich 0,01 ist.

Derartige Verfahren liefern also z.B. Aufschluss darüber, ob eine in der Stichprobe vorhandene Wechselwirkung groß genug ist, um davon auszugehen, dass sie auch in der Zielpopulation vorhanden ist (signifikanter Effekt) oder, ob sie vermutlich lediglich auf zufällige Schwankungen zurückzuführen ist und „in Wahrheit“ kein Trainingseffekt vorliegt (nicht signifikanter Effekt).

3. Bemerkungen zur Stärke signifikanter Effekte

Die Irrtumswahrscheinlichkeit p sinkt jedoch nicht nur mit der Effektgröße, sondern u.a. auch mit steigender Anzahl (N) derjenigen Personen, deren Daten in die Signifikanztestung eingehen. Daher ist es möglich, dass ein Effekt konstanter Stärke in einer großen Stichprobe signifikant wird, in einer kleinen jedoch nicht. Aus diesem Grund ist im Rahmen der Signifikanztestung immer auch eine Angabe zur Stärke des Effekts zu machen, die von der Stichprobengröße unabhängig ist. Eine derartige Angabe erfolgt anhand eines Effektstärkemaßes. Es gibt verschiedene Effektstärkemaße, im Kontext der bei dieser Untersuchung verwendeten Verfahren ist der Koeffizient η^2 (Eta-Quadrat) relevant. Zur Beurteilung der Größe eines Effekts mit η^2 wird konventionell auf folgende Richtgrößen rekuriert:

$\eta^2 < 0,01$ nicht zu beachtender Effekt.

$\eta^2 \geq 0,01 < 0,06$ kleiner Effekt.

$\eta^2 \geq 0,06 < 0,14$ moderater Effekt.

$\eta^2 \geq 0,14$ großer Effekt.

3.2.3.2. Effektivität des Trainings I – Ergebnis der multivariaten Analyse

1. Einführende Bemerkungen

Um den Effekt der im Rahmen des Projekts PiT-Hessen durchgeführten Trainingsmaßnahmen ermitteln zu können, wurde das inferenzstatistische Verfahren der Varianzanalyse herangezogen. Dieses Verfahren ist dazu geeignet, Wechselwirkungen auf Signifikanz zu testen. In der Logik einer Varianzanalyse wird - wie bei den meisten anderen inferenzstatistischen Verfahren - zwischen „unabhängigen Variablen“ (in der Terminologie der Varianzanalyse oft auch als Faktoren bezeichnet²) und „abhängigen Variablen“ unterschieden. Dabei wird davon ausgegangen, dass die unabhängigen Variablen zu einer Veränderung in den abhängigen Variablen führen.

Im Kontext dieser Trainingsevaluation wurden die Variablen Bedingung (Trainingsgruppe vs. Kontrollgruppe) und Messwiederholung (erste Messung vs. zweite Messung) als unabhängige Variablen aufgenommen, die sich auf den Satz der elf Skalen (abhängige Variablen: Konflikteinstieg, Konfliktausstieg usw.) auswirken sollten.

Varianzanalysen, die mehrere unabhängige Variablen berücksichtigen, werden als „mehrfaktoriell“ bezeichnet. Darüber hinaus werden Varianzanalysen als „multivariat“ bezeichnet, wenn sie die Auswirkung der unabhängigen Variablen auf mehrere abhängige Variablen simultan überprüfen, demgegenüber werden Varianzanalysen als „univariat“ bezeichnet, wenn sie die Auswirkung der unabhängigen Variablen auf nur eine abhängige Variable überprüfen. Um die Gefahr von Fehlschlüssen zu minimieren, ist stets zunächst eine multivariate Analyse durchzuführen. Ein signifikanter multivariater Effekt bedeutet, dass die unabhängigen Variablen insgesamt eine signifikante Wirkung ausüben. Weitere Informationen zur Art der Effekte liefert das Verfahren nicht. In diesem Fall ist in der Folge univariat zu überprüfen, in welchen abhängigen Variablen eine signifikante Wirkung vorliegt.

Im Rahmen der Auswertung der Effektivität des PiT-Trainings wurde zunächst eine zweifaktorielle multivariate Varianzanalyse mit Messwiederholung berechnet. Diese Analyse war multivariat, da sie alle elf Skalen als abhängige Variablen einschloss, sie war zweifaktoriell, da sie den Einfluss der beiden Faktoren Bedingung und Zeit sowie deren Wechselwirkung auf die Gruppe der abhängigen Variablen testen sollte und es handelte sich um eine Analyse mit Messwiederholung, da der bereits genannte zweite Faktor ein Messwiederholungsfaktor war. Für einen positiven Trainingseffekt würde eine signifikante multivariate Wechselwirkung der unabhängigen Variablen Bedingung und Zeit auf den Satz der abhängigen Variablen sprechen.

² nicht zu verwechseln mit dem Gebrauch des Begriffs Faktor in der Terminologie der Faktorenanalyse

2. Ergebnis

Die berechnete multivariate Varianzanalyse ergab eine hochsignifikante multivariate Wechselwirkung der Faktoren Bedingung und Messwiederholung auf die Gruppe der elf abhängigen Variablen (Pillai Spur $F(11,807) = 3,81; p = 0,00$). Dieser Effekt ist als klein zu bezeichnen (multivariates $\eta^2 = 0,05$).

Wie bereits geschildert, lässt sich anhand eines derartigen multivariaten Verfahrens jedoch keine Aussage darüber machen, hinsichtlich welcher abhängiger Variablen ein signifikanter Effekt vorliegt. Daher wurde zur Klärung dieser Frage im Anschluss mit jeder der elf abhängigen Variablen eine separate univariate Varianzanalyse berechnet. Durch dieses Verfahren lässt sich ermitteln, ob die beiden Faktoren auf die betreffende abhängige Variable die intendierte signifikante Wechselwirkung ausüben oder nicht. Die Ergebnisse der univariaten Analysen werden im Folgenden geschildert.

3.2.3.3. Effektivität des Trainings II – Ergebnis der univariaten Analysen

1. Allgemeiner Ergebnisüberblick

Insgesamt wurden elf zweifaktorielle univariate Varianzanalysen mit Messwiederholung berechnet. Diese Analyse werden als univariat bezeichnet, da sie jeweils nur eine der elf Skalen als abhängige Variable einschlossen, sie waren zweifaktoriell, da sie den Einfluss der beiden Faktoren Bedingung und Zeit sowie deren Wechselwirkung auf die Gruppe der abhängigen Variablen testen sollten und es handelte sich um Analysen mit Messwiederholung, da der zweite Faktor ein Messwiederholungsfaktor war. Tabelle 8 gibt einen Überblick über die Ergebnisse dieser univariaten Analysen.

Tabelle 8

Skala	Zweifache Wechselwirkung Bedingung *Zeit	inhaltliche Richtung des Effekts	Stärke des Effekts
1. Konflikteinstieg	signifikant	+	klein
2. Konfliktausstieg	signifikant	+	Klein
3. Konformität	nicht sig.		
4. Zivilcourage 1 - Kommunik. mit Umst.	signifikant	+	Klein
5. Zivilcourage 2 - direktes Eingreifen	signifikant	+	klein
6. Selbstwirksamkeit	nicht sig.		
7. Suche nach Lehrerunterstützung	nicht sig.		
8. Angst	nicht sig.		
9. Gewaltbereitschaft	nicht sig.		
10. ausgeübte Aggression	nicht sig.		
11. erfahrene Angriffe	nicht sig.		

Die Tabelle zeigt, dass die beiden Faktoren Bedingung und Messwiederholung auf vier der elf Merkmale eine signifikante Wechselwirkung ausüben. Dabei handelt es sich um die Merkmale Konflikteinstieg, Konfliktausstieg, Zivilcourage 1 - Kommunikation mit Umstehenden sowie Zivilcourage 2 - direktes Eingreifen. Die Inspektion der Richtung der Effekte ergibt, dass alle vier Wechselwirkungen für einen erwünschten Trainingseffekt sprechen (siehe unten). Die Stärke der Effekte ist jeweils als „klein“ zu bezeichnen. Die Effekte könnten jedoch durch mehrere Störfaktoren vermindert worden sein (siehe Fazit zur Ergebnisevaluation).

Im Folgenden werden die Ergebnisse der vier univariaten Analysen mit signifikantem Ergebnis näher geschildert. Die Ergebnisse der übrigen Analysen werden hier nicht weiter berücksichtigt, sie sind jedoch im Anhang dieses Berichts dokumentiert.

2. Skala: Konflikteinstieg

Die univariate Analyse ergab eine hochsignifikante Wechselwirkung der Faktoren Bedingung und Messwiederholung auf die Ausprägung des Merkmals „Fähigkeit zur Vermeidung eines Konflikteinstiegs“ ($F = 6,91$; $p = 0,01$). Der Effekt ist als „klein“ zu bezeichnen (univariates $\eta^2 = 0,01$). Abbildung 6 sowie Tabelle 9 verdeutlichen, wie diese Wechselwirkung zu interpretieren ist:

Abbildung 6

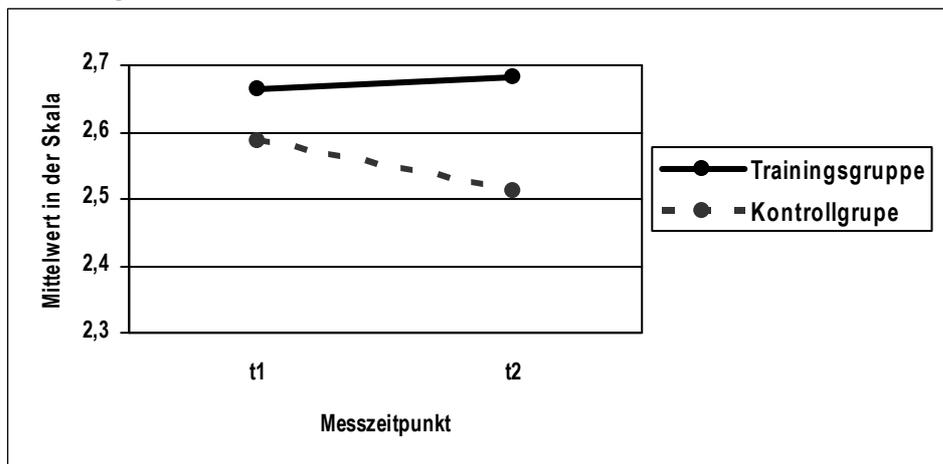


Tabelle 9

Bedingung	Messzeitpunkt		N	p	Effektstärke
	t1	t2			
Trainingsgruppe	2,66 (0,45)	2,68 (0,44)	491	0,01	0,01
Kotrollgruppe	2,59 (0,50)	2,51 (0,54)	328		

In der Tabelle werden unter der Überschrift Messzeitpunkte die Mittelwerte der jeweiligen Bedingungen zum jeweiligen Messzeitpunkt sowie die Standardabweichungen (eingeklammert) aufgeführt. Die Standardabweichung kennzeichnet das Ausmaß der Streuung der in die Berechnung des Mittelwerts eingehenden Einzelwerte. „N“ kennzeichnet die Anzahl der Personen, die unter der jeweiligen Bedingung in die Analyse eingingen. Ebenfalls angegeben werden die Irrtumswahrscheinlichkeit p sowie die Effektstärke η^2 .

In der Trainingsgruppe nimmt die Merkmalsausprägung vom ersten Messzeitpunkt (Mittelwert: 2,66) zum zweiten Messzeitpunkt (Mittelwert: 2,68) geringfügig zu, in der Kontrollgruppe findet demgegenüber über die Zeit eine Reduktion der Ausprägung (Mittelwerte: 2,59 vs. 2,51) statt. Das Datenmuster spricht primär für einen „Puffereffekt“ des Trainings auf die Fähigkeit zur Vermeidung eines Konfliktausstiegs: Es könnte sein, dass das Training gegen eine Verschlechterung der Merkmalsausprägung durch andere Faktoren schützt.

3. Skala: Konfliktausstieg

Die univariate Analyse ergab eine hochsignifikante Wechselwirkung der Faktoren Bedingung und Messwiederholung auf die Ausprägung des Merkmals Fähigkeit zum Konfliktausstieg ($F = 13,61$; $p = 0,00$). Der Effekt ist als „klein“ zu bezeichnen (univariates $\eta^2 = 0,02$). Abbildung 7 sowie Tabelle 10 verdeutlichen, wie diese Wechselwirkung zu interpretieren ist:

Abbildung 7

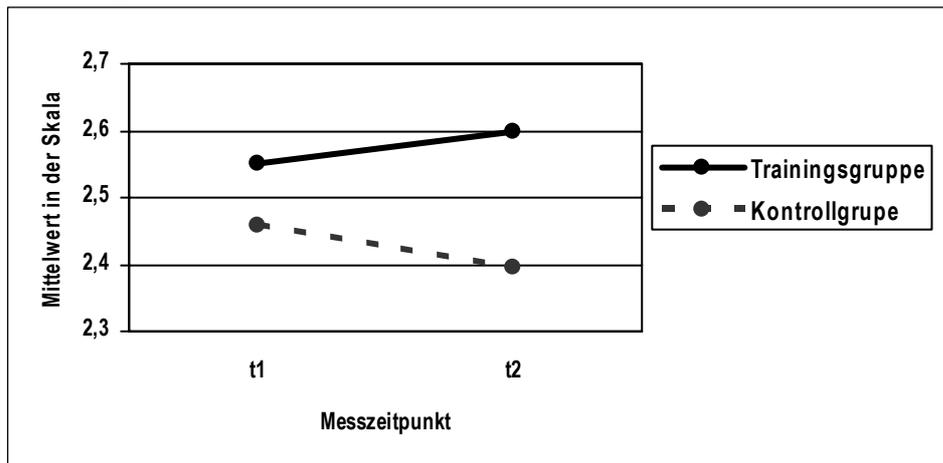


Tabelle 10

Bedingung	Messzeitpunkt		N	p	Effektstärke
	t1	t2			
Trainingsgruppe	2,55 (0,42)	2,60 (0,44)	491	0,00	0,02
Kotrollgruppe	2,46 (0,45)	2,40 (0,44)	328		

Die Merkmalsausprägung nimmt in der Trainingsgruppe vom ersten Messzeitpunkt (Mittelwert: 2,55) zum zweiten Messzeitpunkt (Mittelwert: 2,60) zu, in der Kontrollgruppe findet demgegenüber über die Zeit eine Reduktion der Ausprägung statt (Mittelwerte: 2,46 vs. 2,40), die vom Betrag her nahezu dem Zuwachs in der Trainingsgruppe entspricht. Das Datenmuster spricht für einen erwünschten Effekt des Trainings auf die Fähigkeit zum Konfliktausstieg.

4. Skala: Zivilcourage 1 („Kommunikation mit Umstehenden“)

Die univariate Analyse ergab eine hochsignifikante Wechselwirkung der Faktoren Bedingung und Messwiederholung auf die Ausprägung Merkmals Zivilcourage 1 – Kommunikation mit Umstehenden ($F = 10,99$; $p = 0,00$). Der Effekt ist als „klein“ zu bezeichnen (univariates $\eta^2 = 0,01$). Abbildung 8 sowie Tabelle 11 verdeutlichen, wie diese Wechselwirkung zu interpretieren ist:

Abbildung 8

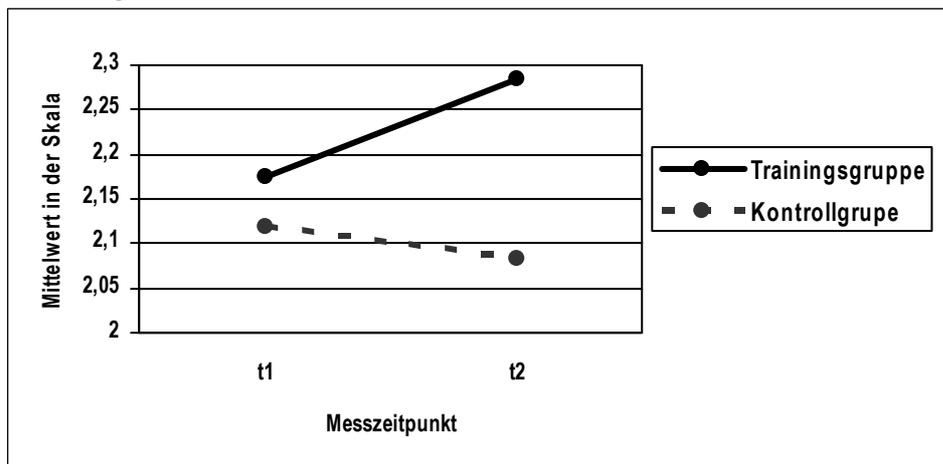


Tabelle 11

Bedingung	Messzeitpunkt		N	p	Effektstärke
	t1	t2			
Trainingsgruppe	2,18 (0,64)	2,28 (0,61)	491	0,00	0,01
Kotrollgruppe	2,12 (0,62)	2,08 (0,66)	328		

In der Trainingsgruppe nimmt die Merkmalsausprägung vom ersten Messzeitpunkt (Mittelwert: 2,18) zum zweiten Messzeitpunkt (Mittelwert: 2,28) zu, in der Kontrollgruppe findet demgegenüber eine Reduktion der Ausprägung (Mittelwerte: 2,12 vs. 2,08) statt. Das Datenmuster spricht für einen erwünschten Effekt des Trainings auf die Ausprägung von Zivilcourage im Sinne der Kommunikation mit umstehenden Personen.

5. Skala: Zivilcourage 2 („direktes Eingreifen“)

Die univariate Analyse ergibt eine hochsignifikante Wechselwirkung der Faktoren Bedingung und Messwiederholung auf die Ausprägung des Merkmals Zivilcourage 2 – direktes Eingreifen ($F = 5,94$; $p = 0,01$). Der Effekt ist als „klein“ zu bezeichnen (univariates $\eta^2 = 0,01$). Abbildung 9 sowie Tabelle 12 verdeutlichen, wie diese Wechselwirkung zu interpretieren ist:

Abbildung 9

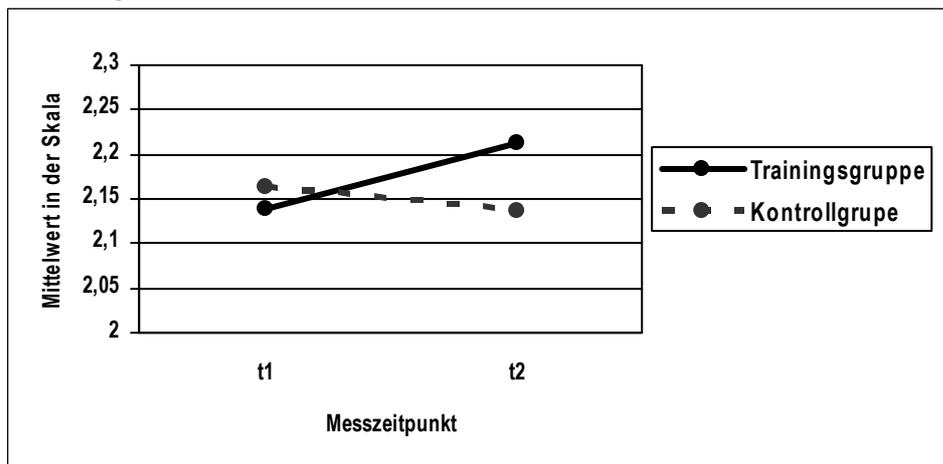


Tabelle 12

Bedingung	Messzeitpunkt		N	p	Effektstärke
	t1	t2			
Trainingsgruppe	2,14 (0,60)	2,21 (0,56)	491	0,02	0,01
Kotrollgruppe	2,16 (0,55)	2,14 (0,56)	328		

In der Trainingsgruppe nimmt die Merkmalsausprägung vom ersten Messzeitpunkt (Mittelwert: 2,14) zum zweiten Messzeitpunkt (Mittelwert: 2,21) zu, in der Kontrollgruppe findet demgegenüber eine minimale Reduktion der Ausprägung (Mittelwerte: 2,16 vs. 2,14) statt. Das Datenmuster spricht für einen erwünschten Effekt des Trainings auf die Ausprägung von Zivilcourage im Sinne von direktem Eingreifen.

3.2.3.4. Effektivität des Trainings III – zum Einfluss weiterer Faktoren

1. Einführung

Im Anschluss an die bereits vorgestellten Auswertungen wurde noch untersucht, ob die dargestellten Ergebnisse zur Effektivität des Trainings noch weiter ausdifferenziert werden müssen. Im diesem Kontext wurde überprüft, welchen Einfluss die folgenden Faktoren auf die Effektivität des Trainings haben:

1. Geschlecht (dichotom: weiblich, männlich).
2. Schultyp (vierstufig: Gymnasium, Realschule, Hauptschule, Gesamtschule).
3. Jahrgangsstufe (dreistufig: 5-6, 7-8, 9).
4. Migrationshintergrund (dichotom: Migrationshintergrund, kein Migrationshintergrund).

Diesbezüglich wäre es beispielsweise denkbar, dass das Training bei SchülerInnen unterschiedlicher Schultypen eine unterschiedliche Wirkung erzielt. Die Überprüfung des Einflusses dieser Faktoren geschah zunächst durch vier dreifaktorielle multivariate Varianzanalysen mit Messwiederholung, in die jeweils einer der oben genannten Faktoren mit aufgenommen wurde. Diese Analysen waren multivariat, da sie alle elf Skalen als abhängige Variablen einschlossen, sie waren dreifaktoriell, da sie den Einfluss der Faktoren Bedingung, Geschlecht / Schultyp / Jahrgangsstufe / Migrationshintergrund und Zeit auf diese abhängigen Variablen überprüfen sollten und es handelte sich um eine Analyse mit Messwiederholung, da der letzte Faktor ein Messwiederholungsfaktor war. Ein Einfluss einer zusätzlich aufgenommen Variablen auf die Trainingswirkung würde sich in einer signifikanten dreifachen Wechselwirkung der drei Faktoren zeigen. Dies würde bedeuten, dass bei der Interpretation eines Trainingseffekts zusätzlich zu berücksichtigen ist, welches Geschlecht, welcher Schultyp, welche Jahrgangsstufe und / oder ob ein Migrationshintergrund vorliegt. Im Folgenden werden die Ergebnisse geschildert.

2. Einfluss des Faktors Geschlecht

Die berechnete multivariate Varianzanalyse ergab eine hochsignifikante dreifache Wechselwirkung der Faktoren Bedingung, Geschlecht und Messwiederholung auf die Gruppe der elf abhängigen Variablen (Pillai Spur $F(11,779) = 2,39$; $p = 0,00$). Der Effekt ist als klein zu bezeichnen (multivariates $\eta^2 = 0,03$). Die nachgeschalteten univariaten Analysen erbrachten jedoch in keiner der elf Skalen eine dreifache Wechselwirkung, die eindeutig im Sinne eines Einfluss des Faktors Geschlecht auf die Trainingswirkung zu interpretieren wäre. Von einer weiteren Darstellung dieser Ergebnisse wird daher abgesehen.

3. Einfluss des Faktors Schultyp

Die berechnete multivariate Varianzanalyse ergab eine hochsignifikante dreifache Wechselwirkung der Faktoren Bedingung, Schultyp und Messwiederholung auf die Gruppe der elf abhängigen Variablen (Pillai Spur $F(33,2325) = 1,71$; $p = 0,00$). Der Effekt ist als klein zu bezeichnen (multivariates $\eta^2 = 0,02$). Hinsichtlich der Ergebnisse der nachgeschalteten Analysen gilt dasselbe wie für den Einfluss des Faktors Geschlecht: In keiner der elf Skalen zeigte sich eine dreifache Wechselwirkung, die eindeutig im Sinne eines Einflusses des Faktors Schultyp auf die Trainingswirkung zu interpretieren wäre.

4. Einfluss des Faktors Jahrgangsstufe

Die berechnete multivariate Varianzanalyse ergab keine signifikante dreifache Wechselwirkung der Faktoren Bedingung, Jahrgangsstufe und Messwiederholung auf die Gruppe der elf abhängigen Variablen (Pillai Spur $F(22,1608) = 1,37$; $p = 0,12$).

5. Einfluss des Faktors Migrationshintergrund

Auch bezüglich dieses Faktors ergab die berechnete multivariate Varianzanalyse keine signifikante dreifache Wechselwirkung der Faktoren Bedingung, Migrationshintergrund und Messwiederholung auf die Gruppe der elf abhängigen Variablen (Pillai Spur $F(11,804) = 0,71$; $p = 0,72$).

3.3. Fazit zur Ergebnisevaluation

Insgesamt zeigt die Ergebnisevaluation ein durchweg positives Bild: In vier der neun direkt zielbezogenen Skalen lässt sich ein positiver Trainingseffekt ermitteln. Dadurch werden fünf der neun Zielvorgaben erfüllt. Dieser Trainingseffekt wird nicht systematisch durch die Faktoren Geschlecht, Schultyp, Jahrgangsstufe und Migrationshintergrund beeinflusst.

Die Effekte sind zwar klein, dies ist aber vor dem Hintergrund potentieller effektmindernder Faktoren zu sehen, wie z.B. der Verwendung eines dreistufigen Antwortformats, das – relativ zu Antwortformaten mit höherer Stufenanzahl – Veränderungen nur eingeschränkt abbilden kann. Darüber hinaus sind die Effekte vor dem Hintergrund einer möglichen Minderung durch die Heterogenität der Trainingsdurchführung zwischen den verschiedenen Projektschulen zu sehen.

Es fällt auf, dass sich die Trainingseffekte lediglich auf der Ebene der Verhaltensintentionen zeigen und zudem in allen dort lokalisierten Skalen vorhanden sind. Auf der Einstellungsebene sowie auf der Ebene der retrospektiven Verhaltensangaben zeigen sich keine Effekte.

Ausgehend von diesen Ergebnissen ist davon auszugehen, dass die SchülerInnen durch das Training primär Verhaltensskripte (Verhaltensrichtlinien) erlernen, die beinhalten, wie sie sich in konkreten konfliktgeladenen Situationen adäquat zu Verhalten haben. Sie lernen, wie sie zu reagieren haben, wenn eine Person einem Konfliktrisiko ausgesetzt ist. Diese Person können sie selbst sein, diesbezüglich verbessert das Training die Fähigkeit zur Vermeidung eines Konflikteinstiegs sowie die Fähigkeit zum Ausstieg aus einem Konflikt. Es kann sich aber auch um eine andere Person handeln, diesbezüglich verbessert das Training das Erkennen der Situation und steigert das Ausmaß des zivilcouragierten Verhaltens.

4. Prozessevaluation

Im Einzelnen werden folgende Unterkapitel werden behandelt:

- 4.1. Prozessevaluation – TeilnehmerInnen.
- 4.2. Prozessevaluation – TeamarbeiterInnen.
- 4.3. Fazit zur Prozessevaluation.

4.1. Prozessevaluation – TeilnehmerInnen

Dieses Unterkapitel ist in die beiden folgenden Abschnitte gegliedert:

- 4.1.1. Methodik: Messinstrument u. Untersuchungsdesign.
- 4.1.2. Auswertungsergebnisse.

4.1.1. Methodik: Messinstrument u. Untersuchungsdesign

Zur Ermöglichung der Prozessevaluation aus der Perspektive der Teilnehmenden wurde ein Fragebogen konstruiert. Das Messinstrument besteht aus zehn Fragen, die erfassen sollen, wie die Teilnehmer das Training hinsichtlich mehrerer Aspekte bewerten. Dabei handelt es sich um acht geschlossene und zwei offene Fragen. Folgende geschlossene Fragen wurden gestellt:

1. Wie interessant fandest du das Training?
2. Konntest du alles verstehen, was besprochen wurde?
3. Hast du durch das Training etwas Neues gelernt?
4. Wie war die Arbeitsatmosphäre?
5. Wie war das Training insgesamt?
6. Traust du dir zu, das Gelernte auch in anderen Situationen zu zeigen?
7. Glaubst du, dass du nach dem Training besser mit Konflikt- und Streitsituationen umgehen kannst?
8. Würdest du einem Freund oder einer Freundin den Besuch solch eines Trainings empfehlen?

Die Fragen werden als „geschlossen“ bezeichnet, da sie mit einem fest vorgegebenem Antwortformat anhand des Ankreuzens einer Zahl zu beantworten sind (siehe auch Messinstrument zur Ergebnisevaluation). Abbildung 10 zeigt exemplarisch einen Ausschnitt aus dem Messinstrument.

Abbildung 10

Wie interessant fandest du das Training?	gar nicht interessant	eher uninteressant	interessant	Sehr interessant
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Konntest du alles verstehen, was besprochen wurde?	nein	eher nein	eher ja	ja
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Die Items sind auf einer vierstufigen Antwortskala zu beantworten. Die exakte Bezeichnung der Antwortkategorien variiert zwischen den Items. Die Angaben lassen sich jedoch itemübergreifend wie folgt interpretieren: Je höher der Wert, desto positiver ist die Trainingsbewertung hinsichtlich des abgefragten Sachverhalts.

An dieser Stelle sei erwähnt, dass die Antwortkategorien im Original-Fragebogen in die andere Richtung kodiert waren, in der Auswertung aber eine Umkodierung in die hier dargestellte Richtung erfolgte. Dies erschien sinnvoll, damit sich die Werte nach dem Schema „je höher, desto besser“ interpretieren lassen.

Weil die Kontrollgruppe nicht am Training teilnahm, stand zur Auswertung keine Vergleichsgruppe zur Verfügung. Die Darstellung und Auswertung erfolgte daher bei den geschlossenen Fragen deskriptiv anhand der Berechnung von Mittelwerten. Zur Mittelwertbildung wurden die Antworten aller Personen auf eine gegebene Frage zunächst aufaddiert und sodann durch die Anzahl eingegangener Personen dividiert.

Im Kontext der beiden offenen Fragen hatten die SchülerInnen die Möglichkeit, ihre Angaben stichwortartig einzutragen. Abbildung 11 zeigt einen Ausschnitt aus dem Messinstrument mit den beiden offenen Fragen.

Abbildung 11

Was hat dir persönlich an dem Training besonders gut gefallen? (bitte ausführen)
Was sollte deiner Meinung nach an dem Training verbessert werden?

Die offene Fragenform hat den Vorteil, dass man auch Informationen über Aspekte erhält, nach denen nicht konkret gefragt wurde, z.B. weil sie nicht bedacht wurden oder bisher nicht eingeplant waren. Diesem Vorteil steht jedoch der Nachteil einer sehr aufwendigen und auch mit größerem subjektivem Interpretationsspielraum behafteten inhaltsanalytischen Auswertung gegenüber. Im Rahmen dieser Auswertung müssen zunächst pro Frage sämtliche Antworten überblickt und sodann übergeordnete Kategorien formuliert werden. Diese Kategorien sollten stets sowohl erschöpfend (alle Antworten lassen sich in eine Kategorie einordnen) als auch möglichst überschneidungsfrei (die Antworten lassen sich möglichst nur einer Kategorie zuordnen) sein.

Durch die Zuordnung der Einzelantworten in die Oberkategorien erhält man eine quantitative Aussage darüber, mit welcher Frequenz in dieser Kategorie geantwortet wurde, d.h. wie häufig die Personen den betreffenden Aspekt aufgeführt haben.

Das vollständige Messinstrument ist im Anhang dokumentiert. Der Fragebogen wurde den Mitgliedern der Trainingsgruppe nach dem Ende der Trainingsdurchführung zusammen mit dem Fragebogen zur Ergebnisevaluation vorgegeben.

4.1.2. Auswertungsergebnisse

Dieses Unterkapitel ist in folgende Abschnitte gegliedert:

- 4.1.2.1. Analysestichprobe.
- 4.1.2.2. Auswertungsergebnis – geschlossene Fragen.
- 4.1.2.2. Auswertungsergebnis – offene Fragen.

4.1.2.1. Analysestichprobe

Die Analysestichprobe besteht aus denjenigen Mitgliedern der Trainingsgruppe, für die zur Ergebnisevaluation zwei verknüpfbare Fragebögen vorlagen.

4.1.2.2. Auswertungsergebnisse – geschlossene Fragen

Tabelle 13 enthält für jedes Item die - über alle Personen der Analysestichprobe - gemittelten Antworten. Diese Mittelwerte können theoretisch zwischen den Polen „1“ (sehr negative Bewertung des Trainings) und „4“ (sehr positive Bewertung des Trainings) lokalisiert sein, es gilt also: Je höher der Mittelwert, desto besser die Bewertung. Der theoretische Skalenmittelwert beträgt 2,5, dieser Wert kennzeichnet eine „neutrale Mitte“, die die Antwortskala in einen Bereich der negativen Bewertung und einen Bereich der positiven Bewertung zerlegt. In der Tabelle werden außerdem die Standardabweichung (Stdabw.) und die Anzahl derjenigen SchülerInnen der Analysestichprobe angegeben, die auf das Item geantwortet haben und somit in die Berechnung des aufgeführten Mittelwertes eingingen (N).

Tabelle 13

	Frage	Mittelwert	Stdabw.	N
1.	Wie interessant fandest du das Training?	2,93	0,71	546
2.	Konntest du alles verstehen, was besprochen wurde?	3,68	0,56	547
3.	Hast du durch das Training etwas Neues gelernt?	3,22	0,98	546
4.	Wie war die Arbeitsatmosphäre?	2,94	0,61	541
5.	Wie war das Training insgesamt?	3,08	0,69	543
6.	Traust du dir zu, das Gelernte auch in anderen Situationen zu zeigen?	3,11	0,87	544
7.	Glaubst du, dass du nach dem Training besser mit Konflikt- und Streitsituationen umgehen kannst?	3,12	0,92	546
8.	Würdest du einem Freund oder einer Freundin den Besuch solch eines Trainings empfehlen?	2,98	1,01	540

Bei der Inspektion der Werte fällt zunächst auf, dass sämtliche Mittelwerte größer als 2,5 sind. Sie übersteigen somit den theoretischen Skalenmittelwert und liegen im „positiven Bewertungsbereich“. Der Mittelwert zur Frage „Konntest du alles verstehen, was besprochen wurde?“ fällt mit 3,68 am positivsten aus. Ebenfalls sehr positiv ausgeprägt ist der Mittelwert zur Frage „Hast du durch das Training etwas gelernt?“ (3,22). Die Mittelwerte zu den Fragen „Würdest du einem Freund oder einer Freundin den Besuch solch eines Trainings empfehlen?“ (2,98), „Wie war die Arbeitsatmosphäre“ (2,94) und „Wie interessant fandest du das Training?“ (2,93) fallen weniger positiv als die Mittelwerte der übrigen Fragen aus, sind aber ebenfalls oberhalb des theoretischen Skalenmittelwerts angesiedelt.

4.1.2.3. Auswertungsergebnisse – offene Fragen

Im Folgenden werden die Ergebnisse der inhaltsanalytischen Auswertung für jede der beiden Fragen dargestellt.

1. Antworten zur Frage: Was hat dir persönlich an dem Training besonders gut gefallen?

Tabelle 14 enthält die Ergebnisse der inhaltsanalytischen Auswertung zur Frage „Was hat dir persönlich an dem Training besonders gut gefallen“. Die Tabelle ist nach Bereichen und Kategorien geordnet. Rechts steht die Anzahl der Personen, die eine Antwort gegeben haben, die in die betreffende Kategorie fällt.

Tabelle 14

Bereich	Kategorie	Anzahl
1. Zentrale Trainingskomponenten	1.1. Rollenspiele / Situationen nachspielen o. vorspielen	306
	1.2. Übungen / Spiele	88
	1.3. Erklärungen / Besprechungen / Diskussionen	30
	1.4. Beteiligung von allen in Gesprächen / Rollenspielen; freie Meinungsäußerung	9
	1.5. Gruppenarbeit	7
	1.6. Einbezug eigener Wünsche / Mitsprache	6
	1.7. Beteiligung versch. Gruppen (Polizei, Jugendarbeit, ...)	5
	1.8. Etwas selber ausprobieren können	4
	Σ 455	
2. Einzelkomponenten des Trainings	2.1. Filme	17
	2.2. Das Ja-und-Nein-Schreien / Schreiübung	17
	2.3. Polizeibesuch	15
	2.4. Rätsel	12
	2.5. Paradoxe Interventionen	6
	2.6. Besuch des jugendlichen Straftäters / "Knasti"	3
	Σ 70	
3. Ergebnis des Trainings	3.1. Gelernt, mit Konfliktsituationen / Gewalt umzugehen	60
	3.2. viel / neues gelernt	28
	Σ 88	
4. Prozess des Trainings	4.1. Spaß gemacht / viel gelacht	19
	4.2. Zusammenarbeit	16
	4.3. Atmosphäre	8
	4.4. Interessant/abwechslungsreich	7
	Σ 50	
5. sonstiges	5.1. alles / fast alles	40
	5.2. nichts / nicht viel	13
	5.3. keine Angabe	60
	5.4. irrelevante Anmerkungen	92

Aus der Tabelle geht hervor, dass die Befragten am häufigsten zentrale Trainingskomponenten ansprechen. Mit besonders hoher Frequenz werden dort Rollenspiele (306 Nennungen; über 50% der Befragten haben in dieser Kategorie geantwortet) sowie Übungen bzw. Spiele im Allgemeinen (88 Nennungen) aufgeführt. Die am dritthäufigsten aufgeführte Inhaltskategorie ist im Bereich der Trainingsergebnisse angesiedelt und lautet: „Gelernt, mit Konfliktsituationen / Gewalt umzugehen“ (60 Nennungen).

Darüber hinaus wird im Bereich „sonstiges“ die Kategorie „alles / fast alles“ 40 mal aufgeführt, demgegenüber die Kategorie „nichts / nicht viel“ nur 13 mal. Insgesamt ist das Ergebnisbild zu dieser Frage als durchweg positiv einzustufen.

2. Antworten zur Frage: Was sollte deiner Meinung nach an dem Training verbessert werden?

Die Ergebnisse der inhaltsanalytischen Auswertung zur Frage; „Was sollte deiner Meinung nach an dem Training verbessert werden?“ sind in Tabelle 15 dargestellt. Der Aufbau der Tabelle entspricht dem Aufbau der vorangehenden Tabelle. Um eine bessere Übersicht zu erhalten, wurde der Bereich „zentrale Trainingskomponenten“ in zwei Unterbereiche zerlegt. Der erste Bereich bezieht sich auf Aspekte, die nach Meinung der Befragten ausgebaut oder intensiviert werden sollten, der zweite Bereich bezieht sich auf Aspekte, die nach Meinung der Befragten reduziert werden sollten.

Tabelle 15

Bereich	Kategorie	Anzahl	
1a) Zentrale Trainingskomponenten – Ausbau	1.1.	abwechslungsreicher (nicht immer das gleiche / Wiederholungen) / interessanter/ lustiger / nicht so langweilig	68
	1.2.	mehr Rollenspiele	32
	1.3.	mehr Übungen / Spiele	29
	1.4.	mehr / längere Pausen (auch zum Nachdenken über das Gelernte)	17
	1.5.	mehr Themengebiete / Situationen	13
	1.6.	realistischer	12
	1.7.	Probleme ausführlicher behandeln, mehr besprechen /erklären, länger diskutieren	9
	1.8.	gesamtes Training länger / häufiger	8
	1.9.	mehr draußen machen	6
	1.10.	mehr auf die Kinder / Kinderwünsche eingehen; Umgang mit den Kindern	6
			Σ 200
1b) Zentrale Trainingskomponenten – Reduktion	1.11.	weniger vorbesprechen / weniger Theorie	41
	1.12.	weniger Übungen / Spiele	26
	1.13.	weniger Rollenspiele / keine Rollenspiele	35
	1.14.	gesamtes Training kürzer	6
	1.15.	kleinere Gruppen	4
		Σ 82	
2. Einzelkomponenten des Trainings	2.1.	mehr Filme	6
	2.2.	keine Rätsel	4
		Σ 10	
3. Ergebnis des Trainings	-	-	-
4. Prozess des Projekts des Trainings	4.1.	Lautstärke / ruhiger / mehr Ernsthaftigkeit / Disziplin (Durchsetzung von Lehrern)	23
	4.2.	Atmosphäre	11
		Σ 34	
5. sonstiges	5.1.	nichts / nicht viel	276
	5.2.	alles / fast alles	50
	5.3.	irrelevante Anmerkungen	105
	5.4.	keine Angabe	90

Der Tabelle ist zu entnehmen, dass die Befragten am häufigsten in der Kategorie „abwechslungsreicher (nicht immer das gleiche / Wiederholungen) / interessanter/lustiger / nicht so langweilig“ antworten (68 Nennungen). Die Kategorie „weniger vorbesprechen / weniger Theorie“ wird 41 mal aufgezählt. Darüber hinaus werden von einigen Befragten mehr Rollenspiele (32 Nennungen) und mehr Übungen bzw. Spiele im Allgemeinen (29 Nennungen) gewünscht, während sich andere Befragte weniger Rollenspiele (35 Nennungen) und Übungen bzw. Spiele im Allgemeinen (26 Nennungen) wünschen. Auffallend ist vor allem auch, dass die Kategorie „nichts / nicht viel“ im Bereich „sonstiges“ 276 mal aufgeführt wird (d.h. über 50% der Befragten haben in dieser Kategorie geantwortet).

4.2. Prozessevaluation – TeamarbeiterInnen

Das Unterkapitel ist in folgende Abschnitte gegliedert:

- 4.2.1. Methodik: Messinstrument u. Untersuchungsdesign.
- 4.2.2. Auswertungsergebnisse.

4.2.1. Methodik: Messinstrument u. Untersuchungsdesign

Das Messinstrument wurde von der AG Sozialpsychologie konstruiert, es besteht aus 15 geschlossenen und drei offenen Fragen und ist im Anhang dokumentiert. Befragt wurden die an den Schulen als Trainer tätigen TeamarbeiterInnen. Die geschlossenen Fragen waren auf einer sechsstufigen Skala zu beantworten. Abbildung 12 zeigt einen Ausschnitt aus dem Messinstrument.

Abbildung 12

5	Wie kam das Training bei den SchülerInnen an?	sehr schlecht	① ② ③ ④ ⑤ ⑥	sehr gut
6	Haben die SchülerInnen im Training etwas zum Thema gelernt?	sehr wenig	① ② ③ ④ ⑤ ⑥	sehr viel

Die genaue Bezeichnung der Antwortkategorien variiert zwischen den Items. Die Antworten lassen sich jedoch stets wie folgt interpretieren: Je höher der Wert, desto positiver ist die Trainingsbewertung hinsichtlich des abgefragten Sachverhalts. Da die Antwortkategorien bereits in der Original-Version in dieser Richtung kodiert waren, musste keine Umkodierung erfolgen.

Im Anschluss an die geschlossenen Fragen wurden die folgenden offenen Fragen gestellt:

1. Wenn es Schwierigkeiten gab: Wo lagen diese?
2. Welche darüber hinaus gehende Unterstützung durch die Leitung des Modellprojekts würden sie sich wünschen?
3. Was würden sie bei erneuter Planung ändern?

4.2.2. Auswertungsergebnisse

4.2.2.1. Auswertungsergebnisse – geschlossene Fragen

Tabelle 16 gibt die Mittelwerte der 15 geschlossenen Fragen wieder. Ebenfalls dargestellt werden Standardabweichung (Stdabw.) und Anzahl der TeilnehmerInnen, die in die Mittelwertberechnung eingingen (N).

Tabelle 16

	Frage	Mittelwert	Stdabw.	N
1.	Wie gut hat die Umsetzung der Trainingsinhalte funktioniert?	4,64	0,75	63
2.	Wie war die Arbeitsatmosphäre während des Trainings?	4,81	0,93	63
3.	Wie hat die Zusammenarbeit im PiT-Team funktioniert?	4,81	1,11	63
4.	Wie kam das Training bei den SchülerInnen an?	4,57	0,89	63
5.	Haben die SchülerInnen im Training etwas zum Thema gelernt?	4,60	0,76	62
6.	Konnten Sie beobachten oder gibt es Rückmeldungen der SchülerInnen, dass sich deren Umgang mit Konflikten in der Praxis verändert hat?	3,83	0,79	52
7.	Für wie notwendig halten Sie es grundsätzlich, dass etwas zur Verbesserung des Konfliktverhaltens von SchülerInnen getan wird?	5,82	0,43	62
8.	Glauben Sie, dass das Training grundsätzlich geeignet ist, Veränderungen im Konfliktverhalten bei den SchülerInnen zu bewirken?	4,94	0,89	64
9.	Für wie groß halten Sie den Bedarf an solchen Konflikttrainings speziell an Ihrer Schule?	5,05	0,85	60
10.	Wie wurde das Training im Kollegium aufgenommen?	4,14	0,96	49
11.	Fühlten Sie sich im Bezug auf die Durchführung u. Organisation des Trainings von der Schulleitung unterstützt?	4,51	1,28	57
12.	Waren Sie mit der Unterstützung durch die Leitung des Modellprojekts (Herr Friedrichs) zufrieden?	4,86	0,89	64
13.	Wie sind Sie rückblickend mit dieser Gestaltung zufrieden?	4,47	1,07	62
14.	War das von Ihnen angesetzte Zeitvolumen angemessen?	3,61	1,00	62
15.	Wie zufrieden waren Sie insgesamt mit dem Training?	4,48	0,84	63

Es fällt zunächst auf, dass sämtliche Mittelwerte größer als der theoretische Skalenmittelwert „3,5“ sind und somit im „positiven Bewertungsbereich“ liegen. Die höchste mittlere Zustimmung erhalten die Fragen: „Für wie notwendig halten Sie es grundsätzlich, dass etwas zur Verbesserung des Konfliktverhaltens von SchülerInnen getan wird?“ (5,82), „Für wie groß halten Sie den Bedarf an solchen Konflikttrainings speziell an Ihrer Schule?“ (5,05), „Glauben Sie, dass das Training grundsätzlich geeignet ist, Veränderungen im Konfliktverhalten bei den SchülerInnen zu bewirken?“ (4,94), „Waren Sie mit der Unterstützung durch die Leitung des Modellprojekts (Herr Friedrichs) zufrieden?“ (4,86) sowie „Wie hat die Zusammenarbeit im PiT-Team funktioniert?“ (4,81). Die positive Reaktion hinsichtlich der letzten Frage gibt erste positive Hinweise zur Evaluation des Teambildungsprozesses. Nur zwei Items haben einen Mittelwert, der kleiner als 4 ist. Dabei handelt es sich um die beiden Fragen: „War das von Ihnen angesetzte Zeitvolumen angemessen?“ (3,61) und „Konnten Sie beobachten oder gibt es Rückmeldungen der SchülerInnen, dass sich deren Umgang mit Konflikten in der Praxis verändert hat?“ (3,83). Aber auch diese Werte liegen im „positiven Bewertungsbereich“.

4.2.2.2. Auswertungsergebnisse – offene Fragen

Im Folgenden werden die inhaltsanalytischen Ergebnisse zu den drei offenen Fragen getrennt aufgeführt. Die einzelnen Äußerungen waren teilweise auf die spezifischen Rahmenbedingungen des jeweiligen Teams und / oder der jeweiligen Berufsgruppe bezogen. Die hier dargestellten Ergebnisse beziehen sich auf Äußerungen, die allgemeinen Aussagecharakter haben.

1. Antworten zur Frage: Wenn es Schwierigkeiten gab: Wo lagen diese?

Tabelle 17 enthält die zentralen Ergebnisse der inhaltsanalytischen Auswertung zur Frage „Wenn es Schwierigkeiten gab: Wo lagen diese?“. Die Tabelle ist nach Bereichen und Kategorien geordnet. Rechts steht die Anzahl der Personen, die eine Antwort gegeben haben, die in die betreffende Kategorie fällt.

Tabelle 17

Bereich		Kategorie		Anzahl
1.	Terminierung und Kommunikation	1.1.	Schwierigkeiten bei Terminplanung und - einholung	10
		1.2.	fehlende / mangelnde / zu geringe Absprachen / Kommunikation zw. den Tea- mern	9
2.	Zu hohe Belastung	2.1.	zu hohe zusätzliche Arbeitsbelastung neben dem Projekt / Überlastung	6
		2.2.	Zeitprobleme / zu wenig Zeit / zu großer Zeit- druck	3
		2.3.	unterschiedliche zeitliche Ressourcen der Teamer	3
3.	Rahmenbedingungen	3.1.	Training auch außerhalb der Schule durchfüh- ren	2
		3.2.	unruhige / unkonzentrierte Schüler	2

2. Antworten zur Frage „Welche darüber hinaus gehende Unterstützung durch die Leitung des Modellprojekts würden Sie sich wünschen?“

Die zentralen Ergebnisse der inhaltsanalytischen Auswertung zur Frage „Welche darüber hinaus gehende Unterstützung durch die Leitung des Modellprojekts würden Sie sich wünschen?“ sind in Tabelle 18 dargestellt.

Tabelle 18

Bereich		Kategorie		Anzahl
1.	Kooperation zw. den Teams	1.1.	inhaltlicher Austausch / Kooperation der Teams untereinander verstärken (z.B. durch regelmäßige Treffen)	8
2.	Unterstützung	2.1.	zusätzliche Unterstützung in Form von mehr Zeit, Personal und / oder Material	7
		2.2.	mehr Fortbildungsmaßnahmen	3
		2.3.	Supervision	2
3.	Planungssicherheit	3.1.	langfristige finanzielle Absicherung gewährleisten	4

3. Antworten zur Frage: „Was würden Sie bei erneuter Planung ändern?“

Tabelle 19 beinhaltet die zentralen Ergebnisse der inhaltsanalytischen Auswertung zur Frage „Was würden Sie bei erneuter Planung ändern?“.

Tabelle 19

Bereich		Kategorie		Anzahl
1.	Verstärkung des Teamaspekts	1.1.	mehr (gemeinsame) Vorbereitungszeit / Planungszeit	12
		1.2.	stärkere Einbindung der einzelnen Berufsgruppen	6
2.	Trainingsdurchführung	2.1.	Zielgruppen des Trainings: Beteiligung aller 7. Klassen des Jahrgangs.	4
		2.2.	Durchführung von Projekttagen / Projektwochen	5
		2.3.	bessere Verankerung im Stundenplan	3
		2.4.	kürzerer Zeitraum / Straffung / kompaktere Gestaltung	3

4.3. Fazit zur Prozessevaluation

Die Prozessevaluation lässt aus der Perspektive der TeilnehmerInnen insgesamt eine positive Bewertung des Trainings erkennen. Der Mittelwert zu jeder geschlossenen Frage liegt im positiven Skalenbereich. Aus der Inhaltsanalyse der offenen Fragen geht hervor, dass die SchülerInnen vor allem die Rollenspiele sowie die Übungen und Spiele im Allgemeinen positiv bewerten. Mehrere SchülerInnen geben darüber hinaus an, dass sie gelernt haben, mit Konfliktsituationen umzugehen. Insgesamt sind die TrainingsteilnehmerInnen zufrieden, dies lässt sich beispielsweise daran erkennen, dass auf die Frage: „Was sollte deiner Meinung nach an dem Training verbessert werden?“ von 276 Personen (dies entspricht über 50% der Befragten) die Kategorie „nichts / nicht viel“ aufgeführt wird. Einige SchülerInnen wünschen sich jedoch eine abwechslungsreichere und interessantere Gestaltung des Trainings.

Aus der Perspektive der TeamarbeiterInnen ergibt sich ebenfalls eine insgesamt positive Bewertung des Trainings. Auch hier liegen die Mittelwerte sämtlicher geschlossener Fragen im „positiven Bewertungsbereich“. Durch die Antworten auf die offenen Fragen kommt zum Ausdruck, dass sich die TeamarbeiterInnen u.a. mehr gemeinsame Vorbereitungszeit, eine stärkere Einbindung der einzelnen Berufsgruppen sowie eine größere Kooperation zwischen den Teams wünschen. Aus den ausformulierten Aussagen (auf den Fragebögen) geht jedoch hervor, dass diese Aspekte als Verbesserungsvorschläge (d.h. das Konzept der Teamarbeit soll noch intensiviert werden) und weniger im Sinne von ernsthaften Schwierigkeiten interpretiert werden sollten.

5. Diskussion

Die zentralen Ergebnisse der Ergebnis- und Prozessevaluation der ersten Trainingswelle des Gewaltpräventionsprojekts „PiT-Hessen (Prävention im Team)“ lassen sich wie folgt zusammenfassen:

- (1) Die Prozessevaluation zeigt – sowohl aus der Perspektive der TrainingsteilnehmerInnen als auch aus der Perspektive der TeamarbeiterInnen – eine insgesamt positive Bewertung des Projekts „PiT-Hessen“.
- (2) Die Daten der Ergebnisevaluation sprechen dafür, dass durch das Training fünf der neun Trainingsziele erreicht werden. Dabei handelt es sich um die Ziele: „Entwicklung persönlicher Handlungsoptionen“, „Hinsehen statt Wegesehen“, „Stärkung von Zivilcourage“, „Erlernen von „richtigem“ Helferverhalten in gewaltbesetzten Situationen“ sowie „Frühzeitiges Einbinden von „Helfern“ in Problemsituationen“.
- (3) Die Effekte sind klein, lassen sich aber zufallskritisch absichern und sind zudem vor dem Hintergrund möglicher effektmindernder Faktoren zu sehen (z.B. eingeschränkte Möglichkeit zur Anzeige von Veränderungen aufgrund des dreistufigen Antwortformats sowie Heterogenität der Trainingsdurchführung).
- (4) Die Trainingseffekte sind nur im Bereich der Verhaltensintentionen in spezifischen konfliktgeladenen Situationen zu beobachten und z.B. nicht auf der Einstellungsebene.
- (5) Daher ist davon auszugehen, dass die TeilnehmerInnen durch das Training Skripte (Richtlinien) dafür erlernen, wie sie sich in spezifischen konfliktgeladenen Situationen zu Verhalten haben. Sie lernen, wie sie zu reagieren haben, wenn eine Person in „Not“ ist:
 - a. Hinsichtlich der eigenen Person lernen die TeilnehmerInnen, wie sie den Einstieg in einen drohenden Konflikt vermeiden und wie sie aus einem bereits entstandenen Konflikt aussteigen können.
 - b. Zudem lernen sie, anderen Personen, die Opfer eines physischen oder psychischen Angriffs durch einen Aggressor sind, zu helfen.

- (6) Vor diesem Hintergrund erscheint die Anwendung und Weiterentwicklung des Projekts „PiT“ Hessen aus der wissenschaftlichen Perspektive sinnvoll.
- (7) Als Evaluationsperspektiven für zukünftige Evaluationen des Projekts PiT-Hessen bieten sich – neben der bereits in Planung befindlichen Evaluation der Ausstrahlungswirkung der Trainingsmaßnahmen auf die Schule – u.a. folgende Möglichkeiten an:
- a. Untersuchung der Langzeitwirkung der Trainingsmaßnahmen durch Einsatz des vorhandenen Messinstruments zu einem dritten Messzeitpunkt.
 - b. Untersuchung des Effekts der Trainingsmaßnahmen auf das tatsächliche Verhalten, z.B. durch Anwendung von Verhaltensbeobachtungs- und beurteilungsverfahren.
 - c. Untersuchung des Effekts der Trainingsmaßnahmen an anderen Zielgruppen.

6. Literaturverzeichnis

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.

Heitmeyer, W., Collmann, B., Conrads, J., Matuschek, I., Kraul, D., Kühnel, W., Möller, R. & Ulbrich-Herrmann, M. (1995). *Gewalt: Schattenseiten der Individualisierung in verschiedenen sozialen Milieus*. Weinheim: Juventa.

Labuhn, A.S., Wagner, U., van Dick, R. & Christ, O. (2004). Determinanten zivilcouragierten Verhaltens: Ergebnisse einer Fragebogenstudie. *Zeitschrift für Sozialpsychologie*, 35, 93-103.

Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.

Schwarzer, R. (1993). *Measurement of perceived self-efficacy. Psychometric scales for cross-cultural research*. Berlin: Freie Universität Berlin.

Sturzbecher, D. (1997). *Jugend und Gewalt in Ostdeutschland. Lebenserfahrungen in Schule, Freizeit und Familie*. Göttingen: Verlag für angewandte Psychologie

Wagner, U. (2004). Ansätze und Ergebnisse von Projektevaluation: Einige Überlegungen zu Präventionsprogrammen gegen Fremdenfeindlichkeit. *Journal für Konflikt- und Gewaltforschung*, 6, 8-18.

7. Anhang³

A 1 Messinstrument – Ergebnisevaluation

A 2 Messinstrument – Prozessevaluation (TeilnehmerInnen)

A 3 Messinstrument – Prozessevaluation (TeamarbeiterInnen)

A 4 Ergebnisse der univariaten Varianzanalysen

³ Weitere Dokumente befinden sich im Anhang des Zwischenberichts zur Evaluation der ersten Trainingswelle des Gewaltpräventionsprojekts „PiT-Hessen (Prävention im Team)“



Fragebogen zum Konfliktverhalten

Liebe Schülerin, lieber Schüler!

Nach dem Ende des Trainings PiT sind wir nun noch einmal daran interessiert zu erfahren, wie du dich in Konfliktsituationen verhältst. Deshalb bitten wir dich, wieder einige Fragen zu beantworten.

Es ist sichergestellt, dass deine Antworten völlig anonym bleiben, d.h. niemand wird erfahren, wie du geantwortet hast.

Kreuze bitte die Fragen an der Stelle an, die deiner Meinung oder deinem Verhalten am ehesten entspricht.

Beispiel:

	ja	manchmal	nein
Ich gehe gerne ins Kino.	1	2	3

Wenn du gerne ins Kino gehst, müsstest du die 1 (ja) ankreuzen.

Wenn du nur manchmal gerne ins Kino gehst, müsstest du die Ziffer 2 ankreuzen. Dies ist im Beispiel der Fall.

Wenn du nicht gerne ins Kino gehst, müsstest du die 3 (nein) ankreuzen.

Gibt's noch Fragen? Wenn ja, wende dich bitte an deine Interviewerin oder deinen Lehrer.

Los geht's!

Was glaubst du, wie gefährlich könnten folgende Situationen für dich werden?

Auf dem Weg zur S-Bahn willst du dir am Kiosk noch schnell etwas zu trinken kaufen. An einigen Stellen stehen Gruppen von Fußballfans herum, die von einem Spiel kommen, manche halten Bierdosen in den Händen. Als du am Kiosk stehst, ruft einer dir zu: „He Kleine(r) komm mal rüber und gib einen aus!“

	gefährlich	weiß nicht	ungefährlich
Was glaubst du, wie gefährlich könnte die Situation für dich werden?	1	2	3

Du liegst auf einer Wiese im Park in der Sonne. In einiger Entfernung spielen mehrere Typen Fußball. Sie sind etwas älter als du. Plötzlich landet der Ball vor deinen Füßen. Einer der Jugendlichen kommt auf dich zugelaufen.

	gefährlich	weiß nicht	ungefährlich
Was glaubst du, wie gefährlich könnte die Situation für dich werden?	1	2	3

Im Kaufhaus steht auf der Rolltreppe plötzlich ein unbekannter Typ neben dir, er ist ungefähr in deinem Alter und quatscht dich an. Erst fragt er dich nach der Uhrzeit und dann, ob er mal mit deinem Handy einen Freund anrufen könne, wegen einer Party, sein Handy sei gerade leer. Du könntest auch mitkommen auf die Party. Dir kommt der Typ seltsam und ziemlich aufdringlich vor.

	gefährlich	weiß nicht	ungefährlich
Was glaubst du, wie gefährlich könnte die Situation für dich werden?	1	2	3

Du bist mit einer Freundin verabredet. Als Treffpunkt habt ihr den Springbrunnen auf dem Marktplatz ausgemacht. Du wartest schon eine Weile. Auf der anderen Seite des Brunnens hängen zwei Typen rum, etwas älter als du. Sie machen Sprüche über die Vorübergehenden. Plötzlich merkst du, dass sie anfangen über dich zu reden. Du kannst nicht hören, was sie sagen, aber es scheint, als würden sie sich über dich lustig machen. Jetzt ruft dir der eine etwas zu, was du aber nicht verstehen kannst.

	gefährlich	weiß nicht	ungefährlich
Was glaubst du, wie gefährlich könnte die Situation für dich werden?	1	2	3

Im Folgenden möchten wir von dir wissen, wie du dich in bestimmten Situationen verhältst.

Versuche zunächst, dich in die folgende Situation hineinzusetzen:

Du warst mit einer Gruppe von Freunden im Kino. Anschließend wollt ihr mit dem Bus nach Hause fahren. Als ihr an der Bushaltestelle wartet, hält ein Auto an. Ein paar ältere Typen sitzen drin, die einen deiner Freunde kennen. Sie halten Bierdosen in den Händen und scheinen nicht mehr ganz nüchtern zu sein. Sie bieten an, euch zu einer Party mitzunehmen. Deine Freunde sind gleich Feuer und Flamme und wollen mitfahren.

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu :	ja	vielleicht	nein
Ich steige ein und denke, es wird schon nichts passieren.	1	2	3
Ich versuche, meine Freunde zu überzeugen, dass es zu gefährlich ist, bei den Typen mitzufahren.	1	2	3
Ich steige <u>nicht</u> ein und warte lieber auf den nächsten Bus.	1	2	3

Nun eine andere Situation:

Es ist Nachmittag und du bist auf dem Weg zu deiner besten Freundin oder deinem besten Freund. Vom Bus aus musst du noch ein Stück laufen. An der nächsten Ecke siehst du drei Typen stehen, etwas älter als du, die da gelangweilt herumhängen. Sie sehen nicht besonders nett aus und du denkst dir, Mist, wenn das mal keinen Ärger gibt.

Als du bei ihnen vorbeikommst ruft einer von ihnen: „Hey du, komm mal her!“

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu:	ja	vielleicht	nein
Ich gehe mit großem Abstand an den Typen vorbei, ohne sie zu beachten.	1	2	3
Ich gehe dicht an den Typen vorbei und gucke sie dabei an, um ihnen zu zeigen, dass ich mich nicht einschüchtern lasse.	1	2	3
Ich gehe zu den Typen hin um zu sehen, was sie von mir wollen.	1	2	3
Ich bleibe mit großem Abstand zu den Typen stehen, um mit ihnen zu reden.	1	2	3
Ich gehe an den Typen vorbei und beschimpfe sie dabei.	1	2	3

Stell dir vor:

Es ist Pause in der Schule und du stehst mit einigen Freundinnen und Freunden aus der Klasse in einer Ecke des Schulhofs. Euch ist langweilig. Als Timo vorbeigeht, fangen die anderen an, über ihn zu lästern. Timo geht in die Parallelklasse und ist schon zweimal sitzen geblieben. Er ist etwas langsam, deshalb wird er oft geärgert, aber du findest ihn eigentlich ganz nett. Einer deiner Freunde sagt, „Los, den schnappen wir uns heute nach der Schule“. Du weißt, was das bedeutet: Die anderen wollen Timo verprügeln.

Wie verhältst du dich?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu:	ja	vielleicht	nein
Ich sage den anderen, dass ich dabei <u>nicht</u> mitmache.	1	2	3
Ich sage lieber nichts dazu, weil ich mich bei den anderen nicht unbeliebt machen möchte.	1	2	3
Ich mache mit.	1	2	3
Ich mache mit, beschließe aber, Timo heimlich zu warnen.	1	2	3
Ich versuche, die anderen zu überzeugen, dass Timo eigentlich ganz nett ist und es nicht in Ordnung ist, ihn zu verprügeln.	1	2	3
Ich erfinde eine Ausrede, dass ich nach der Schule nicht dabei sein kann.	1	2	3

Stell dir vor:

Du sitzt auf einer Bank im Park, hörst Musik von deinem MP3-Player und wartest auf einen Freund, mit dem du hier verabredet bist. Ein dir unbekannter Typ, etwas älter als du, kommt angeschlendert und setzt sich direkt neben dich. Er legt dir den Arm um die Schulter und meint: „Los Kleine/r lass mal deinen MP3-Player rüberwachsen, sonst passiert hier gleich was.“

Wie verhältst du dich?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu:	ja	vielleicht	nein
Ich tue so, als ob nichts passiert wäre.	1	2	3
Ich hoffe, dass er von selbst aufhört.	1	2	3
Ich versuche, den Typen davon zu überzeugen, mich in Ruhe zu lassen.	1	2	3
Ich zeige deutlich, dass ich nicht einverstanden bin.	1	2	3
Ich beleidige oder bedrohe den Typen.	1	2	3
Ich bitte Leute, die vorbeigehen, mir zu helfen.	1	2	3
Ich greife den Typen körperlich an.	1	2	3
Ich versuche, Vorbeigehende aufmerksam zu machen, indem ich laut werde.	1	2	3
Ich stehe auf und gehe weg.	1	2	3
Ich diskutiere in ruhigem Ton mit dem Typen darüber, dass ich ihm meinen MP3-Player nicht geben möchte.	1	2	3

Stell dir vor:

Du bist mit zwei Freunden in der Stadt. Ihr wollt ein Geburtstagsgeschenk für eine Freundin besorgen, jeder soll 5,-€ dazugeben. Ihr habt an eine CD gedacht, findet aber keine, die ihr gefallen könnte. Die Geschäfte machen gleich zu und der Geburtstag ist am selben Abend. Ihr sucht weiter und findet schließlich eine coole Lampe, die sie sich auch wünscht, die aber 45,- € kosten soll. Das wären 15,-€ für jeden. Alle sind abgenervt von der Sucherei und deine Freunde wollen, dass ihr die Lampe kauft. 15,-€ sind für dich viel Geld, du hast in diesem Monat kaum noch Taschengeld und deine Eltern geben dir kein Geld für Geschenke. Außerdem hast du auch noch Schulden vom letzten Monat.

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu:	ja	vielleicht	nein
Ich sage, dass ich mit dem Kauf der Lampe einverstanden bin und hoffe, dass ich das Geld irgendwie beschaffen kann.	1	2	3
Ich sage, dass ich mit dem Kauf der Lampe nicht einverstanden bin, weil ich nicht so viel Geld habe.	1	2	3
Ich versuche, die anderen zu überzeugen, dass wir bestimmt noch etwas anderes finden werden.	1	2	3

Stell dir vor:

Es ist Nachmittag und du hängst mit ein paar Freundinnen und Freunden zu Hause rum. Ihr verbringt eure freie Zeit meistens zusammen, seid eine richtig gute Truppe. Daniel, der neu in eurer Klasse ist, möchte auch noch bei euch mitmachen. Er ist o.k., soll aber erst mal zeigen, dass er das auch verdient hat. Boris schlägt vor, dass ihr raus in den alten Steinbruch fahrt, dort gibt es eine steile Felswand, die soll Daniel hochklettern, als Mutprobe. Du kennst den Steinbruch auch. Es ist ziemlich gefährlich, dort zu klettern, überall brechen die Steine weg, man kann leicht abstürzen. Die anderen finden die Idee gut.

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu:	ja	vielleicht	nein
Ich sage, dass ich die Idee mit der Mutprobe an der Felswand gut finde.	1	2	3
Ich versuche, die anderen davon zu überzeugen, dass das Klettern an der Felswand zu gefährlich ist.	1	2	3
Ich sage nichts und hoffe, dass die anderen sich für eine andere Mutprobe entscheiden.	1	2	3
Ich sage, dass ich <u>nicht</u> damit einverstanden bin, dass Daniel die Felswand hochklettern soll.	1	2	3

Und noch zwei weitere Situationen:

Du beobachtest, wie an der Bushaltestelle zwei Mädchen von einem Typen, etwas älter als du, beschimpft werden. Sie schauen verängstigt.

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu	ja	vielleicht	nein
Ich tue so, als ob ich die Mädchen kennen würde, verwickle sie in ein Gespräch und gehe mit ihnen weg.	1	2	3
Ich gehe lieber weiter.	1	2	3
Ich fordere den Typen ruhig aber bestimmt auf, die Mädchen in Ruhe zu lassen.	1	2	3
Ich warte erst mal ab, ob andere Leute eingreifen.	1	2	3
Ich versuche, andere Leute anzusprechen, um sie zum Eingreifen zu bringen.	1	2	3
Ich rufe den umstehenden Leuten zu: „Guckt nicht zu, helft lieber!“	1	2	3
Ich frage die Mädchen, was los ist.	1	2	3
Ich fordere die Leute, die sich in meiner Nähe befinden auf, mit mir gemeinsam etwas gegen das Verhalten des Typen zu tun.	1	2	3
Ich ziehe die Mädchen von dem Jungen weg.	1	2	3

Du beobachtest im Bus, wie mehrere Typen, etwas älter als du, einen Jungen, so alt wie du, beschimpfen, bespucken und rumschubsen.

Wie verhältst du dich in dieser Situation?

Kreuze bitte bei jedem Verhalten an, was auf dich zutrifft:

Dieses Verhalten trifft auf mich zu	ja	vielleicht	nein
Ich tue so, als ob ich den Jungen kennen würde, verwickle ihn in ein Gespräch und gehe mit ihm weg.	1	2	3
Ich gehe lieber weiter.	1	2	3
Ich fordere die Typen ruhig aber bestimmt auf, den Jungen in Ruhe zu lassen.	1	2	3
Ich warte erst mal ab, ob andere Leute eingreifen.	1	2	3
Ich versuche, andere Leute anzusprechen, um sie zum Eingreifen zu bringen.	1	2	3
Ich rufe den umstehenden Leuten zu: „Guckt nicht zu, helft lieber!“	1	2	3
Ich spreche den Jungen an und frage, was los ist.	1	2	3
Ich fordere die Leute, die sich in meiner Nähe befinden auf, mit mir gemeinsam etwas gegen das Verhalten der Typen zu tun.	1	2	3
Ich ziehe den Jungen von den Typen weg.	1	2	3

Eine ernste Streitsituation oder eine Situation, in der man sich bedroht fühlt, kann dazu führen, dass man Angst hat.

Was denkst du dazu?

	ja	weiß nicht	nein
Ich finde, Angst zu haben ist dumm.	1	2	3
Ich glaube, Angst kann mir helfen zu erkennen, wie gefährlich eine Situation ist.	1	2	3
Ich finde, wer Angst hat, macht sich lächerlich.	1	2	3
Wenn ich Angst habe, versuche ich sie sofort zu unterdrücken, denn Angst ist uncool.	1	2	3

Was glaubst du, wie du allgemein mit bedrohlichen Situationen oder Streit umgehen kannst?

	ja	weiß nicht	nein
Ich komme auch mit überraschenden Streitsituationen klar.	1	2	3
Wenn ich mich anstrengte, kann ich auch in bedrohlichen Situationen immer eine Lösung finden.	1	2	3
In bedrohlichen Situationen weiß ich immer, wie ich mich verhalten soll.	1	2	3
Wenn eine bedrohliche Situation auftaucht, werde ich schon damit fertig.	1	2	3

Im Folgenden geht es um deine Einstellung zu gewalttätigem Verhalten.

	ja	weiß nicht	nein
Um mich durchzusetzen, muss ich manchmal auch Gewalt anwenden.	1	2	3
Ich finde es gut, wenn es Leute gibt, die mit Gewalt für Ordnung sorgen.	1	2	3
Wenn mich jemand beleidigt, kann es sein, dass er sich eine fängt.	1	2	3
Ich finde es schlimm, wenn Leute sich prügeln.	1	2	3
Manchmal muss ich Gewalt einsetzen, um nicht den Kürzeren zu ziehen.	1	2	3
Wenn es wirklich drauf ankommt, bin ich auch bereit, Gewalt anzuwenden.	1	2	3
Mit Gewalt schafft man klare Verhältnisse.	1	2	3

Nun interessiert uns noch, ob du Lehrerinnen oder Lehrer um Hilfe bittest, wenn es ernsthaften Streit gibt.

Bei ernsthaftem Streit in der Schule...

	ja	vielleicht	nein
...würde ich mich hilfesuchend an eine Lehrerin oder einen Lehrer wenden.	1	2	3
...glaube ich, dass eine Lehrerin oder ein Lehrer helfen kann, den Streit zu schlichten.	1	2	3
...halte ich die Lehrerinnen und Lehrer lieber raus.	1	2	3

Zuletzt möchten wir noch von dir wissen:

Wie oft hast du in den letzten zwei Wochen...

	kein mal	1-2 mal	3-5 mal	mehr als 5 mal
...andere geschlagen oder getreten?	1	2	3	4
...andere gezwickt?	1	2	3	4
...andere an den Haaren gezogen?	1	2	3	4
...andere angespuckt?	1	2	3	4

Wie oft bist du in den letzten zwei Wochen von anderen ...

	kein mal	1-2 mal	3-5 mal	mehr als 5 mal
...mit Worten angegriffen worden?	1	2	3	4
...geschlagen oder getreten worden?	1	2	3	4

Wie alt bist du? _____ Jahre

Wie heißt deine Schule? _____

Welche Klasse besuchst du? _____

Auf was für eine Schule gehst du?

Hauptschule Realschule Gymnasium

Gesamtschule Sonderschule Förderstufe

Dein Geschlecht ist weiblich männlich

Was ist deine Nationalität? _____

In welchem Land wurdest du geboren? _____

In welchem Land wurde deine Mutter geboren? _____

In welchem Land wurde dein Vater geboren? _____

Ist Deutsch deine Muttersprache? ja nein

Trage bitte noch folgendes Kürzel ein, damit bei einer weiteren Befragung die Fragebögen einander zugeordnet werden können.

Dein eigener Geburtstag

z.B. für den 07. Mai:

0	7
S	a

Die ersten beiden Buchstaben
deines Vornamens

z.B. für „Sandra“:

Zunächst möchten wir wissen, wie dir das Training PiT gefallen hat. Deshalb bitten wir dich, die folgenden Fragen zu beantworten. Es kommt uns dabei nur auf deine Meinung an, es gibt also keine richtigen oder falschen Antworten.

Wie interessant fandest du das Training?	sehr interessant <input type="radio"/>	interessant <input type="radio"/>	eher uninteressant <input type="radio"/>	gar nicht interessant <input type="radio"/>
Konntest du alles verstehen, was besprochen wurde?	ja <input type="radio"/>	eher ja <input type="radio"/>	eher nein <input type="radio"/>	nein <input type="radio"/>
Hast du durch das Training etwas neues gelernt?	ja <input type="radio"/>	eher ja <input type="radio"/>	eher nein <input type="radio"/>	nein <input type="radio"/>
Wie war die Arbeitsatmosphäre?	sehr gut <input type="radio"/>	ziemlich gut <input type="radio"/>	ziemlich schlecht <input type="radio"/>	sehr schlecht <input type="radio"/>
Wie war das Training insgesamt?	sehr gut <input type="radio"/>	ziemlich gut <input type="radio"/>	ziemlich schlecht <input type="radio"/>	sehr schlecht <input type="radio"/>
Traust du dir zu, das Gelernte auch in anderen Situationen zu zeigen?	ja <input type="radio"/>	eher ja <input type="radio"/>	eher nein <input type="radio"/>	nein <input type="radio"/>
Glaubst du, dass du nach dem Training besser mit Konflikt- und Streitsituationen umgehen kannst?	ja <input type="radio"/>	eher ja <input type="radio"/>	eher nein <input type="radio"/>	nein <input type="radio"/>
Würdest du einem Freund oder einer Freundin den Besuch solch eines Trainings empfehlen?	ja <input type="radio"/>	eher ja <input type="radio"/>	eher nein <input type="radio"/>	nein <input type="radio"/>

**Was hat dir persönlich an dem Training besonders gut gefallen?
(bitte ausführen)**

Was sollte deiner Meinung nach an dem Training verbessert werden?

Liebe Mitglieder der PiT-Teams,

nach Abschluss des Trainings ist uns Ihre Einschätzung des Trainingsverlaufs wichtig. Wir bitten Sie, die nachstehenden Fragen zu beantworten.

1	Wie gut hat die Umsetzung der Trainingsinhalte funktioniert?	sehr schlecht	① ② ③ ④ ⑤ ⑥	sehr gut
2	Wie war die Arbeitsatmosphäre während des Trainings?	sehr schlecht	① ② ③ ④ ⑤ ⑥	sehr gut
3	Wie hat die Zusammenarbeit im PiT-Team funktioniert?	sehr schlecht	① ② ③ ④ ⑤ ⑥	sehr gut

4	Wenn es Schwierigkeiten gab: Wo lagen diese?			
---	--	--	--	--

5	Wie kam das Training bei den SchülerInnen an?	sehr schlecht	① ② ③ ④ ⑤ ⑥	sehr gut
6	Haben die SchülerInnen im Training etwas zum Thema gelernt?	sehr wenig	① ② ③ ④ ⑤ ⑥	sehr viel
7	Konnten Sie beobachten oder gibt es Rückmeldungen der SchülerInnen, dass sich deren Umgang mit Konflikten in der Praxis verändert hat?	Starke Verschlechterung des Konfliktverhaltens	① ② ③ ④ ⑤ ⑥	Starke Verbesserung des Konfliktverhaltens
8	Für wie notwendig halten Sie es grundsätzlich, dass etwas zur Verbesserung des Konfliktverhaltens von SchülerInnen getan wird?	gar nicht notwendig	① ② ③ ④ ⑤ ⑥	sehr notwendig
9	Glauben Sie, dass das Training grundsätzlich geeignet ist, Veränderungen im Konfliktverhalten bei den SchülerInnen zu bewirken?	gar nicht geeignet	① ② ③ ④ ⑤ ⑥	sehr gut geeignet
10	Für wie groß halten Sie den Bedarf an solchen Konflikttrainings speziell an Ihrer Schule?	kein Bedarf	① ② ③ ④ ⑤ ⑥	sehr großer Bedarf
11	Wie wurde das Training im Kollegium aufgenommen?	sehr ablehnend	① ② ③ ④ ⑤ ⑥	sehr aufgenommen
12	Fühlten Sie sich im Bezug auf die Durchführung und Organisation des Trainings von der Schulleitung unterstützt?	gar nicht unterstützt	① ② ③ ④ ⑤ ⑥	sehr gut unterstützt
13	Waren Sie mit der Unterstützung durch die Leitung des Modellprojekts (Herr Friedrichs) zufrieden?	gar nicht zufrieden	① ② ③ ④ ⑤ ⑥	sehr zufrieden

14	Welche darüber hinaus gehende Unterstützung durch die Leitung des Modellprojekts würden Sie sich wünschen?
----	--

15	Wie haben Sie die Durchführung des Trainings über den Verlauf des Schuljahres zeitlich gestaltet? (Mehrfache Nennungen sind möglich)	Projektwoche 0	Einzelne Projektstage 0	Einzelne Schul- stunden 0
----	---	-----------------------	-----------------------------------	--

16	Wie sind Sie rückblickend mit dieser Gestaltung zufrieden?	gar nicht zufrieden	① ② ③ ④ ⑤ ⑥	sehr zufrieden
17	War das von Ihnen angesetzte Zeitvolumen angemessen?	zu wenig Zeit	① ② ③ ④ ⑤ ⑥	zu viel Zeit
18	Wie zufrieden waren Sie insgesamt mit dem Training?	gar nicht zufrieden	① ② ③ ④ ⑤ ⑥	sehr zufrieden

19	Was würden Sie bei erneuter Planung ändern?
----	---

20	Welcher Berufsgruppe gehören Sie an?	LehrerIn 0	Polizei 0	Jugendhilfe 0
----	--------------------------------------	-------------------	------------------	----------------------

21	Wie heißt Ihre Schule?
----	------------------------

Vielen Dank!

Skala	Beding.	Messzeitpunkte		N	p	Effektstärke
		t1	t2			
1. Konflikteinstieg	TG	2,66 (0,45)	2,68 (0,44)	491 328	0,01	0,01
	KG	2,59 (0,50)	2,51 (0,54)			
2. Konfliktausstieg	TG	2,55 (0,42)	2,60 (0,44)	491 328	0,00	0,02
	KG	2,46 (0,45)	2,40 (0,44)			
3. Konformität	TG	2,74 (0,36)	2,68 (0,38)	491 328	0,28	0,00
	KG	2,69 (0,38)	2,67 (0,40)			
4. Zivilcourage 1 („Kommunikation“)	TG	2,18 (0,64)	2,28 (0,61)	491 328	0,00	0,01
	KG	2,12 (0,62)	2,08 (0,66)			
5. Zivilcourage 2 („Eingreifen“)	TG	2,14 (0,60)	2,21 (0,56)	491 328	0,02	0,01
	KG	2,16 (0,55)	2,14 (0,56)			
6. Selbstwirksamkeit	TG	2,22 (0,49)	2,30 (0,45)	491 328	0,22	0,00
	KG	2,32 (0,45)	2,37 (0,46)			
7. Suche nach Lehrerunterstützung	TG	2,37 (0,63)	2,35 (0,61)	491 328	0,63	0,00
	KG	2,26 (0,64)	2,21 (0,67)			
8. Angst	TG	2,59 (0,63)	2,63 (0,47)	491 328	0,96	0,00
	KG	2,55 (0,64)	2,59 (0,50)			
9. Gewaltbereitschaft	TG	1,97 (0,61)	1,97 (0,97)	491 328	0,08	0,00
	KG	2,00 (0,64)	2,07 (0,65)			
10. Ausgeübte Aggression	TG	3,65 (0,46)	3,54 (0,66)	491 328	0,57	0,00
	KG	3,57 (0,57)	3,48 (0,59)			
11. Erfahrene Angriffe	TG	3,20 (0,83)	3,08 (0,91)	491 328	0,94	0,00
	KG	3,20 (0,83)	3,07 (0,86)			