

**Die Kunst, wirksames Feedback zu
formulieren: Empirische Gegenüberstellung
der Konzepte *gut vs. schlecht* und
beibehalten vs. verändern beim Geben und
Nehmen von Feedback**

**The Art of Phrasing Effective Feedback: Empirical Comparison
of the Concepts *Good vs. Bad* and *Maintaining vs. Changing* in
Providing and Taking Feedback**

Masterarbeit

in Psychologie (M. Sc.)

Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik,
Personal- und Sozialpsychologie

Otto-Friedrich-Universität Bamberg

Autorin: Anna-Katharina Christa Kunigunda Stöcker

Betreuerin: Prof. Dr. Astrid Schütz



Bamberg, 2022

Dieses Werk ist als freie Onlineversion über das Forschungsinformationssystem (FIS; <https://fis.uni-bamberg.de>) der Universität Bamberg erreichbar. Das Werk steht unter der CC-Lizenz CC-BY.

Lizenzvertrag: Creative Commons Namensnennung 4.0
<http://creativecommons.org/licenses/by/4.0>.



URN: <urn:nbn:de:bvb:473-irb-554135>
DOI: <https://doi.org/10.20378/irb-55413>

Vorwort

Die vorliegende Masterarbeit vergleicht das neue Konzept der Feedbackformulierung *beibehalten vs. verändern* mit der gängigen Formulierungsvariante *gut vs. schlecht*. Die Arbeit stellt eine Prüfungsleistung im Masterstudiengang Psychologie (M. Sc.) an der Otto-Friedrich-Universität Bamberg dar.

Ein Dank sei an dieser Stelle gerichtet an Frau Prof. Dr. Astrid Schütz für ihre Unterstützung bei der Verwirklichung des Projektes und für ihre (wortwörtlich) allzeit hilfreiche Betreuung, an das Team des gesamten Lehrstuhls für die theoretischen und praktischen Hilfestellungen, an Dr. Friederike S. Bornträger für die gemeinsame Entwicklung der Idee und den Austausch, mit dem sie mich in besonderem Maße fordert und fördert, an Mara Münder für ihre Begleitung über die Entstehungsphase dieser Arbeit hinweg, und an meine Familie für die Ermöglichung dieses Studiums und von vielem darüber hinaus.

Vor Beginn der eigentlichen Arbeit noch ein Wort zur geschlechtergerechten Sprache: In dieser Arbeit wird bewusst auf die Verwendung des generischen Maskulinums verzichtet, da dies zu einer geringeren gedanklichen Einbeziehung von Frauen als von Männern führt (Irmen & Linner, 2005). Aus Lesbarkeitsgründen wird – wenn möglich – auf geschlechtsneutrale Bezeichnungen zurückgegriffen (Irmen & Linner, 2005). Wenn eine solche Formulierung nicht möglich ist, wird mit Hilfe des sogenannten Gender-Sternchens das Kontinuum und die Vielfalt der Geschlechtlichkeit (im Gegensatz zum bipolaren Geschlechterkonzept) hervorgehoben – wohlwissend, dass dies derzeit nicht dem Regelwerk der Amtlichen deutschen Rechtschreibung entspricht (Rat für deutsche Rechtschreibung, 2018).

Anna-Katharina Stöcker

Bamberg, den 01. September 2018

Kurzzusammenfassung

Feedback dient dem Menschen als soziales Wesen dazu, seine subjektive Perspektive zu erweitern und ist ein Thema von enormer Relevanz. Gleichzeitig liegen divergierende Befunde zur Wirksamkeit vor. Diese Diskrepanz aufzuklären, ist Ziel dieser Arbeit. Im Fokus steht die Formulierung von Feedback. In Studie 1 zum Feedbackgeben wurden Rückmeldungen von 78 Studierenden zur Lehrevaluation von Universitätsveranstaltungen untersucht. Hypothesen waren, dass Feedback formuliert als *beibehalten vs. verändern* annehmbarere, spezifischere und prozessbezogenere, also stärker verhaltens- und weniger eigenschaftsbezogene Rückmeldungen generiert, als die Formulierung *gut vs. schlecht*. In Studie 2 zum Feedbacknehmen wurde untersucht, ob nach *beibehalten vs. verändern* formuliertes Feedback zu Assessment-Center-Aufgaben für die 70 Feedbacknehmenden annehmbarer, hilfreicher und verhaltensänderungsrelevanter ist. Explorativ wurde in Studie 2 die Selbstwertschätzung einbezogen. Die Teilnehmenden gaben (Studie 1) bzw. erhielten (Studie 2) dabei im between-subject-Design das Feedback entweder in der Variante *beibehalten vs. verändern* oder *gut vs. schlecht*. Beim Geben von Feedback nach *beibehalten vs. verändern* werden signifikant weniger eigenschaftsbezogene Inhalte genannt. Das Erhalten von Feedback nach *beibehalten vs. verändern* führte zu einer signifikant positiveren Einstellung gegenüber dem Feedback. Nach Feedback generell und in der *gut vs. schlecht*-Bedingung fassten Personen mit niedriger Selbstwertschätzung signifikant mehr Vorsätze als Personen mit hoher Selbstwertschätzung. In der *beibehalten vs. verändern*-Bedingung fassten hingegen Personen mit hoher Selbstwertschätzung mehr Vorsätze. Folglich hat die Formulierung von Feedback als *beibehalten vs. verändern* für Feedbackgebende wie -nehmende positive Effekte. Weitere Studien sind durchzuführen, um zusätzliche Unterschiede zwischen den Konzepten herauszuarbeiten und den Ansatz möglicherweise als Bewältigung der paradoxen Effekte von Feedback zu etablieren.

Abstract

Feedback serves humans as social beings to broaden their subjective perspective and is a topic of enormous relevance. Simultaneously, there are divergent findings on efficacy. Clarifying this discrepancy is the aim of this work. The focus is how to phrase feedback. In Study 1 (providing feedback), feedback on university courses generated by 78 students was examined. Hypotheses were that feedback phrased as *maintaining vs. changing* was more acceptable, more specific, and more process-related, i.e., more behavioral-, and less trait-related, than feedback phrased as *good vs. bad*. In study 2 (taking feedback), it was examined whether feedback phrased as *maintaining vs. changing* on assessment center tasks was more acceptable, helpful and behavior change-relevant for the 70 feedback takers. Exploratively, self-esteem was included in the analysis. The participants provided (study 1) or took (study 2) feedback in a between-subject-design either phrased as *maintaining vs. changing* or *good vs. bad*. Feedback phrased as *maintaining vs. changing* was significantly less trait-related. Taking feedback phrased as *maintaining vs. changing* resulted in a significantly more positive attitude towards feedback. After taking feedback, people with low self-esteem show significantly more intentions to change their behavior than persons with high self-esteem, whereas in the *maintaining vs. changing* group people with high self-esteem have more intentions. Consequently, feedback phrased as *maintaining vs. changing* effects both feedback providing and taking in a positive way. Further studies are needed to identify other differences between the concepts and ultimately to establish the approach as addressing the paradoxical effects of feedback.

Inhaltsverzeichnis

Abbildungsverzeichnis	7
Tabellenverzeichnis	8
Abkürzungsverzeichnis	9
1 Einführung	10
2 Theorie und Stand der Forschung	11
2.1 Feedback	11
2.1.1 Definition und Arten	11
2.1.2 Wirksamkeit von Feedback	12
2.1.3 Wörtliche Formulierung von Feedback	16
2.2 Statisches und dynamisches Denkkonzept	17
2.2.1 Grundlagen	17
2.2.2 Feedback und Denkkonzepte	19
2.3 Theoretisches Modell dieser Arbeit und Hypothesen	20
3 Studie 1: Feedbackgeben	23
3.1 Methode	23
3.1.1 Untersuchungsdesign und -ablauf	23
3.1.2 Stichprobe	24
3.1.3 Fragebögen	26
3.1.4 Kodierung	26
3.2 Ergebnisse	28
3.2.1 Einstellung zum Feedback	28
3.2.2 Anzahl der genannten Aspekte	29
3.2.3 Konkretheit des Feedbacks	30
3.2.4 Verhaltensbezogenheit der Feedbackinhalte	31
3.2.5 Eigenschaftsbezogenheit der Feedbackinhalte	32
3.3 Diskussion	33
4 Studie 2: Feedbacknehmen	37
4.1 Methode	37
4.1.1 Untersuchungsdesign	37

4.1.2	Untersuchungsablauf	38
4.1.3	Stichprobe	40
4.1.4	Fragebögen	43
4.1.5	Kodierung	44
4.2	Ergebnisse	45
4.2.1	Einstellung zum Feedback	45
4.2.2	Nützlichkeit des Feedbacks	47
4.2.3	Konkretheit der Vorsätze	49
4.2.4	Anzahl der Vorsätze	50
4.2.5	Explorative Untersuchungen zur Selbstwertschätzung	51
4.3	Diskussion	53
4.3.1	Diskussion der Hypothesen	53
4.3.2	Diskussion zu Selbstwertschätzung und Feedback	56
5	Gesamtdiskussion	58
	Literaturverzeichnis	60
	Anhang	68
	Eidesstattliche Versicherung	95

Abbildungsverzeichnis

<i>Abbildung 1.</i> Wirkfaktoren von Feedback.....	15
<i>Abbildung 2.</i> Darstellung des Zusammenhangs zwischen Denkkonzept und Verhalten nach Dweck (1986).	18
<i>Abbildung 3.</i> Hypothesen zum Feedbackgeben (Studie 1).....	21
<i>Abbildung 4.</i> Hypothesen zum Feedbacknehmen (Studie 2).....	21
<i>Abbildung 5.</i> Zeitlicher Ablauf in Studie 2.....	39
<i>Abbildung 6.</i> Prozentuale Häufigkeit der Ratingangaben von 1 = <i>überhaupt nicht</i> bis 7 = <i>voll und ganz</i>	49
<i>Abbildung 7.</i> Anzahl der Vorsätze nach Selbstwertschätzungs-Gruppe und Feedbackvariante.	53

Tabellenverzeichnis

Tabelle 1. <i>Mittelwert und Standardabweichung der Einstellung zum Feedbackgeben generell, im aktuellen Fall und als Differenz.....</i>	29
Tabelle 2. <i>Mittelwert und Standardabweichung der Einstellung zum Feedbacknehmen generell, im aktuellen Fall und als Differenz.....</i>	47

Abkürzungsverzeichnis

BV-Bedingung	Bedingung, in der die Teilnehmenden Feedback in der Formulierung <i>beibehalten</i> vs. <i>verändern</i> gegeben bzw. erhalten haben
GS-Bedingung	Bedingung, in der die Teilnehmenden Feedback in der Formulierung <i>gut</i> vs. <i>schlecht</i> gegeben bzw. erhalten haben

1 Einführung

Alle Menschen erhalten tagtäglich Feedback von anderen (Donnelly & Kirk, 2015) – ob in formellen Gesprächen mit Mitarbeitenden oder Assessment-Center-Teilnehmenden, in Form von informellen Äußerungen wie „Gut gemacht!“ oder nur das Lächeln des Kassierers an der Supermarktkasse. Da der Mensch ein soziales Wesen ist (für einen Überblick aus soziologischer Perspektive vgl. Vester, 2009) und jeweils nur aus seiner subjektiven und damit natürlich eingeschränkten Perspektive die Welt wahrnehmen kann, braucht er den Austausch mit und die Rückmeldung von anderen (vgl. shared reality-Theorie; für einen Überblick siehe Echterhoff, Higgins & Levine, 2009; Echterhoff, 2012). Neben diesem sogenannten extrinsischen Feedback ist im weiter gefassten Sinne aber auch intrinsisches Feedback möglich, z. B. das Gefühl, etwas gut gemacht zu haben oder besser machen zu können (Donnelly & Kirk, 2015).

Gerade im beruflichen Kontext wird man sich zunehmend bewusster, welche Rolle Feedback spielt: Viele Unternehmen schmücken sich mit ihrer Feedbackkultur. Wie diese jedoch im Unternehmen tatsächlich gelebt wird, ist unterschiedlich. Gefordert wird Feedback als Standard im Rahmen von Assessment-Centern (Arbeitskreis Assessment Center e.V., 2016) und belegt ist zumindest der Zusammenhang von Feedback mit dem Karriereaufstieg (Dimotakis, Mitchell & Maurer, 2017). Wie die Feedbackkultur von Feedbacknehmenden erlebt wird, ist von der von Unternehmen dargestellten Feedbackkultur abweichend. So klagen Angestellte über ein regelrechtes Feedbackvakuum – man erhalte, wenn überhaupt, nur Feedback, wenn man aktiv danach suche (Ashford, Blatt & VandeWalle, 2003). Zum Teil werden Feedbackprozesse gar als katastrophal oder Zeitverschwendung beschrieben (Bouskila-Yam & Kluger, 2011). Aber nicht nur im Arbeitskontext, sondern auch im Erziehungs- und Bildungskontext und im Feld der Persönlichkeitsentwicklung spielt Feedback eine zentrale Rolle (Hattie, 1999; Hattie & Timperley, 2007; McKimm, 2009).

Feedback ist für das soziale Wesen Mensch wichtig, wird in der Praxis als Instrument genutzt und von Personen aktiv eingefordert – dennoch gibt es divergierende Befunde bezüglich seiner Wirksamkeit (vgl. Kapitel 2.1.2.1). Ein Grund für diese Diskrepanz könnte in der Formulierung von Feedback liegen: Es gibt verschiedene Formulierungsansätze, die in bisherigen Studien noch nicht systematisch untersucht worden sind. Die vorliegende Masterarbeit beinhaltet erste Studien zur Thematik.

2 Theorie und Stand der Forschung

2.1 Feedback

2.1.1 Definition und Arten

Der Begriff *Feedback* kommt ursprünglich aus der Kybernetik und umfasst den Abgleich von Ist- und Sollwerten (Burkert, 2017). Eingang in die Psychologie fand der Begriff durch Modelle wie das Selbstregulationsmodell von Carver und Scheier (1981), in dem der Begriff einen Rückmeldeprozess in einem Regelkreis beschreibt. Feedback dient demnach dazu, durchgeführte Handlungen mit Zielvorstellungen abzugleichen und gegebenenfalls durch zieldienlichere Handlungsalternativen zu ersetzen.

Feedback gilt als die Rückmeldung von Information an ein Individuum oder eine Gruppe (Besieux, 2017) mit dem Ziel der Veränderung von Verhalten (McKimm, 2009). Dabei hat sich die Unterscheidung zwischen *Person-*, *Prozess-* und *Performanzfeedback* etabliert (Brummelman, Thomaes, Overbeek et al., 2014; Kamins & Dweck, 1999). Personfeedback umfasst Rückmeldungen über Eigenschaften und Fähigkeiten, weshalb auch von eigenschaftsbezogenem Feedback gesprochen wird. Prozessfeedback bezieht sich auf das Verhalten der Person, im Speziellen auf ihre Anstrengung und die verwendeten Strategien, weshalb auch von verhaltensbezogenem Feedback gesprochen wird (Brummelman, Thomaes, Overbeek et al., 2014; Kamins & Dweck, 1999). Performanzfeedback spiegelt die Aufgabenbewältigung, also Leistung und Ergebnis, wider (Kamins & Dweck, 1999).

Je nach Studie werden die Arten von Feedback weiter differenziert oder anders benannt. Hattie und Timperley (2007) unterteilen das von Kamins und Dweck (1999) als Prozessfeedback zusammengefasste Feedback weiter nach dessen Zweck: Aufgabenverarbeitungs- und Selbstregulationsfeedback. Corpus, Ogle und Love-Geiger (2006) unterscheiden beispielsweise Sozialvergleichslob und Bewältigungslob, was dem Bereich Personfeedback respektive dem Prozessfeedback zuzuordnen ist.

Die unterschiedliche Differenzierung und Benennung erschwert den Vergleich verschiedener Studien. Diese Unschärfe erklärt die Diversität der Wirksamkeitsbefunde, auf die in Kapitel 2.1.2 noch näher eingegangen wird, zumindest teilweise.

Im Fokus der meisten Studien steht Performanz-, seltener Prozessfeedback (Gabelica, van den Bossche, Segers & Gijssels, 2012), wobei der inhaltliche Schwerpunkt der Literatur

überwiegend im Bildungskontext liegt: Zahlreiche Studien beschäftigen sich mit der Leistungsrückmeldung in Schule oder Universität (Carless, 2006; Evans, 2013; McKimm, 2009).

2.1.2 Wirksamkeit von Feedback

2.1.2.1 Paradoxa der Feedbackwirkung

Widersprüchlichkeit von Effektrichtungen und -stärken. Die Wirkung von Feedback ist in zahlreichen Studien untersucht und in Metaanalysen zusammengefasst worden. Je nach Metaanalyse zeigen sich geringfügige (Cohen, 1980), aber auch mittlere bis starke Effekte hinsichtlich der Leistungsverbesserung (Kluger & DeNisi, 1996). Hervorzuheben ist, dass in über einem Drittel der Primärstudien in der Metaanalyse von Kluger und DeNisi (1996) sogar negative Effekte von Performanzfeedback gefunden worden sind. Auch wenn dieser negative Zusammenhang in anderem Kontext, z. B. Teamfeedback, nicht auftritt (Gabelica et al., 2012), ist die große Varianz der Ergebnisse – gerade im Kontrast zur in Kapitel 1 beschriebenen hohen praktischen Relevanz von Feedback – bemerkenswert (Hattie & Timperley, 2007; Villiers, 2013).

Brinko (1993) beschreibt Feedback infolge der Effektvarianz als mächtiges, aber nicht konsistent erfolgreiches Instrument.

Unterschiedliche Wirkrichtungen je nach Art des Lobs. Je nachdem, ob Person- oder Prozesslob formuliert wird, zeigen sich unterschiedliche Wirkrichtungen. Vor allem Personlob scheint negativen Konsequenzen zur Folge zu haben:

- Durch sozial vergleichendes Lob (Personlob) wird die intrinsische Motivation bei Kindern gedämpft, wohingegen sie mittels Lob für Aufgabenbewältigung (Prozesslob) wird intrinsische Motivation bei Kindern gestärkt wird (Corpus et al., 2006).
- Personlob führt bei Kindern mit geringer Selbstwertschätzung bei späterem Misserfolg zu Scham, sie attribuieren die Ursache dafür auf sich selbst als Person, bewerten sich selbst schlechter und weisen einen negativen Affekt auf (Brummelman, Thomaes, Overbeek et al., 2014; Cimpian, Arce, Markman & Dweck, 2007; Kamins & Dweck, 1999; Mueller & Dweck, 1998).

Die negativen Auswirkungen von Personlob könnten damit zusammenhängen, dass selbiges den Fokus auf die Person an sich lenkt und Misserfolge folglich auf das Selbst attribuiert werden (Brummelman, Thomaes, Overbeek et al., 2014). Als Alternativerklärung für die widersprüchlichen Befunde schlagen Kamins und Dweck (1999) das Konzept der bedingten Aufmerksamkeit bzw. Wertschätzung vor: Das Kind wird nur bei Erfolg als Person wertgeschätzt, was zwar einerseits zur Ausführung des gewünschten Verhaltens führen kann, aber auch diverse negative Konsequenzen wie Scham und Schwankungen in der Selbstwertschätzung nach Misserfolgen hat (Assor, Roth & Deci, 2004).

Auch wenn die Gründe für die negativen Konsequenzen von Personlob bislang nicht eindeutig geklärt sind, ist unstrittig, dass Feedback prozessbezogen sein muss, also auf veränderbare Verhaltensweisen abzielen muss, um Veränderungen hervorbringen zu können (Donnelly & Kirk, 2015; Evans, 2013; Gibbs & Simpson, 2004; Pomerantz & Kempner, 2013).

Unterschiedliche Wirkweise je nach Spezifität des Feedbacks. Die Spezifität von Feedback, also der konkrete Verhaltensbezug in der Formulierung, scheint ebenso mit der Wirkrichtung von Feedback in Zusammenhang zu stehen. So zeigen sich innerhalb des Prozesslobs verschiedene Wirkrichtungen von Feedback: Bei Kindern mit geringer Selbstwertschätzung führt sogenanntes überhöhtes Prozesslob (z. B. „Du hast eine unglaublich schöne Zeichnung erstellt!“ im Gegensatz zu spezifischen, verhaltensbezogenen Rückmeldungen) zu einer Vermeidung von Herausforderungen (Brummelman, Thomaes, Orobio de Castro, Overbeek & Bushman, 2014). Darüber hinaus können vor allem auf Basis spezifischen Feedbacks Verhaltensvorsätze getroffen werden (van der Leeuw et al., 2016), die in der Konsequenz zu tatsächlichen Verhaltensänderungen führen (Aguinis et al., 2012; Gollwitzer & Sheeran, 2006). Dies ist in der Regel (für Gegenbefunde siehe destruktives Feedback in Kapitel 2.1.2.3) Ziel von Feedback (McKimm, 2009).

Unterschiedliche Wirkweise von Feedback je nach Höhe der Selbstwertschätzung. Nicht nur von Seiten der Feedbackgebenden, auch auf Seiten der Feedbacknehmenden kann die Wirkung von Feedback moduliert werden, z. B. durch die Höhe der Selbstwertschätzung der*des Feedbacknehmenden. Negatives Feedback führt bei Personen mit geringer Selbstwertschätzung zu negativen Erwartungen bezüglich erfolgreicher Aufgabenbewältigung (Shrauger & Rosenberg, 1970), die Personen sind demotivierter

nach negativem Feedback, erleben emotionalen Stress, fühlen sich unwohler und zeigen geringere Performanz in Folgeaufgaben (im Vergleich zu positivem Feedback und zu Personen mit hoher Selbstwertschätzung; Brockner, 1979; Brockner et al., 1983; Brown & Marshall, 2001; Brown, 2010; Shrauger & Sorman, 1977). In einer naturalistischen Feldstudie zeigte sich ein stärkerer Performanzrückgang nach negativem Feedback bei Personen mit niedriger Selbstwertschätzung (Brockner, Derr & Laing, 1987).

Forschung zeigt: Nicht jedes Feedback zeigt positive Wirkung, nicht jedes Lob verstärkt erwünschtes Verhalten, nicht alle Personen profitieren von der gleichen Art von Feedback gleichermaßen.

2.1.2.2 Wirkfaktoren von Feedback

Um die Varianz hinsichtlich der Wirksamkeit von Feedback zu erklären, werden in dieser Arbeit die Wirkfaktoren genauer beleuchtet. In Abbildung 1 werden genannte Wirkfaktoren verschiedener Veröffentlichungen aufgeführt.

Die Literatur diesbezüglich ist divers – sowohl was die Anzahl und Spezifität der genannten Wirkfaktoren betrifft, als auch bezüglich der empirischen Qualität und Aussagekraft. Während Donnelly und Kirk (2015) sowie Gibbs und Simpson (2004) jeweils einen pragmatischen Leitfaden zum Feedbackgeben im Kontext der medizinischen Fachausbildung bzw. von Universitäten bereitstellen, sind die Arbeiten von Villiers (2013) und Evans (2013) als wissenschaftliche Literaturreviews einzuordnen. Auch unter Berücksichtigung dieser sich unterscheidenden Zielgruppen ist die Diversität der Kriterien effektiven Feedbacks auffällig – einzig das Kriterium der zeitnahen Rückmeldung wird in allen Studien genannt (Donnelly & Kirk, 2015; Evans, 2013; Gibbs & Simpson, 2004; Villiers, 2013), die übrigen Kriterien werden unterschiedlich häufig aufgegriffen.

Kriterien	Studie			
	Donnelly und Kirk (2015)	Gibbs und Simpson (2004)	Villiers (2013)	Evans (2013)
<i>Inhalt</i>				
Performanz- und/oder Prozessfeedback, kein Personfeedback				
Spezifität/ Detaillierung				
Zielorientierung				
inkl. Verbesserungsvorschläge				
bedeutsame Hinweise				
Anpassung an Situation				
Anpassung an Verständnis der Schüler*innen				
Beginn mit Lob				
<i>Organisatorischer Rahmen</i>				
zeitnahe Rückmeldung				
umsetzbare Hinweise				
regelmäßige Durchführung				
in vertraulicher Situation				
behutsames Vorgehen				
Integration in den Lernprozess				
explizite Kriterienorientierung				
Feedback auf noch nicht Fertigstelltes				
Anregung zur Selbstreflexion				
Selbstwertförderung				
Wahlfreiheit bei Leistungserhebung				
Feedback zur Verbesserung des Lehrprozesses				
<i>Aktive Rolle der*s Feedbacknehmenden</i>				
aufmerksames Annehmen				
Einbindung der Schüler*innen				
Dialog statt Monolog				
Feedback zur Vernetzung zwischen Lehrkräften und Schüler*innen				
Peerfeedback				
Training für Peerfeedback und Selbstbewertung				

Abbildung 1. Wirkfaktoren von Feedback. Bei Grünfärbung ist das jeweilige Kriterium in der Quelle vermerkt, bei Rotfärbung nicht.

2.1.2.3 Annahmen über Gründe zur Diversität der Befunde

Zu den paradoxen Effektrichtungen sowie den sich unterscheidenden Kriterien wirksamen Feedbacks werden an dieser Stelle verschiedene Annahmen diskutiert. Eine Ursachenklärung fehlt nach dem aktuellen Stand der Forschung und ist möglicherweise in der Formulierung von Feedback zu finden.

Eine grundlegende Problematik ist die fehlende Definition und vorherrschende Breite an verwendeten Begriffen, wie bereits in Kapitel 2.1.1 diskutiert (Besieux, 2017). Dies erschwert die Erstellung von Reviews und Metaanalysen.

Des Weiteren ist zu berücksichtigen, dass Feedback nicht nur unwirksam, sondern auch destruktiv sein kann (Besieux, 2017). Gerade schwächenfokussierte Rückmeldungen zeigen weitere negative Effekte zusätzlich zu ausbleibender Leistungssteigerung (Aguinis et al., 2012). Neben Unzufriedenheit von Mitarbeitenden kann solches Feedback auch Abwehrreaktionen und Ablehnung der Inhalte bewirken (Fedor, Eder & Buckley, 1989; Fedor, Rensvold & Adams, 1992; Ilgen, Fisher & Taylor, 1979), insbesondere wenn es nicht rücksichtsvoll vorgetragen wird (Steelman & Rutkowski, 2004). Personfeedback führt zu hilflosen und selbstbeschuldigenden Reaktionen sowie, selbst wenn es positiv ist, zu Verwundbarkeit (Kamins & Dweck, 1999). Bei Zusammenfassung all dieser Befunde in Metaanalysen überrascht es nicht, dass es je nach inhaltlichen Ein- und Ausschlusskriterien zu stark variierenden Effektstärken kommt.

Eine letzte Annahme bezieht sich auf die Methodik: Die Wirksamkeit von Feedback präzise zu messen, ist laut Price, Handley, Millar und O'Donovan (2010) schwierig, vielleicht sogar unmöglich. Problematische Aspekte sind u. a. der Spielraum in der Interpretation des Feedbacks durch den Feedbacknehmenden und die zusätzlichen Wirkfaktoren, die neben dem Feedback ebenso eine Verhaltensänderung anstoßen könnten (Price et al., 2010). Weitere Probleme der Messung von Feedbackwirksamkeit werden im Methoden- und Diskussions- teil der vorliegenden Arbeit (vgl. Kapitel 3.1, 3.3, 4.1 und 4.3) diskutiert.

2.1.3 Wörtliche Formulierung von Feedback

Ein Charakteristikum von Feedback ist die verwendete Formulierung. Diese ist bislang wenig erforscht. Dabei deuten Arbeiten wie die von Steelman und Rutkowski (2004) darauf hin, dass die Formulierung eine aufschlussreiche Quelle der Varianz von Effektrichtungen

sein könnte. Eine in der Praxis viel propagierte Formulierungsvariante – das sogenannte „Feedbacksandwich“, also im Feedback zuerst etwas positives, dann negatives und zum Schluss wieder etwas positives zurückzumelden – ist jedoch widerlegt: Versuchsteilnehmende präferieren die Reihenfolge negativ, positiv, positiv (Slowiak & Lakowske, 2017).

Einige Studien untersuchten schriftliches Feedback hinsichtlich jeweils verschiedener Wirkfaktoren (vgl. Kapitel 2.1.2.2). Dabei zeigten sich selbst bei Auszubildenden im medizinischen Bereich – Personen, die häufig Feedback geben – Tendenzen in der Formulierung wie geringe Spezifität, personbezogenes statt prozess- oder performanzbezogenes Feedback und fehlende Handlungsimplicationen (Canavan, Holtman, Richmond & Katsufakis, 2010). Ähnliche Ergebnisse wurden auch bei Peerfeedback unter Allgemeinärzt*innen gefunden (Prins, Sluijsmans & Kirschner, 2006).

Sluijsmans, Brand-Gruwel und van Merriënboer (2002) konnten zeigen, dass Feedbackgeben trainierbar ist: nach einem Training gaben trainierte Peers stärker spezifisches, konstruktiveres und strukturierteres Feedback. Dies unterstreicht eine Veränderbarkeit der Formulierung von Feedback nach entsprechender Sensibilisierung.

2.2 Statisches und dynamisches Denkkonzept

2.2.1 Grundlagen

Die Forschungsgruppe um Carol Dweck hat eine Theorie zum statischen und dynamischen Denkkonzept entwickelt. Ihre Arbeit liefert diverse Anknüpfungspunkte zur Wirkweise von Feedback, die in Kapitel 2.2.2 näher dargestellt werden. Zunächst sollen jedoch die Grundlagen des Modells dargestellt werden.

Dweck und Kollegen unterscheiden zwischen einem *statischen* und einem *dynamischen* Konzept des Denkens, der Intelligenz und des Selbst (im englischen Original *fixed vs. growth mindset*). Die Studien von Dweck und Kolleg*innen wurden im Wesentlichen mit Kindern durchgeführt, die Denkkonzepte und deren Konsequenzen lassen sich aber auch an Erwachsenen nachweisen (Dweck, Chiu & Hong, 1995b, 1995a; Dweck, 2002; Rhodewalt, 1994).

Aus dem jeweiligen Denkkonzept von Personen lässt sich Verhalten in Leistungssituationen ableiten, wie genau stellt das Grundgerüst des Modells in Abbildung 2 dar (nach Dweck, 1986).

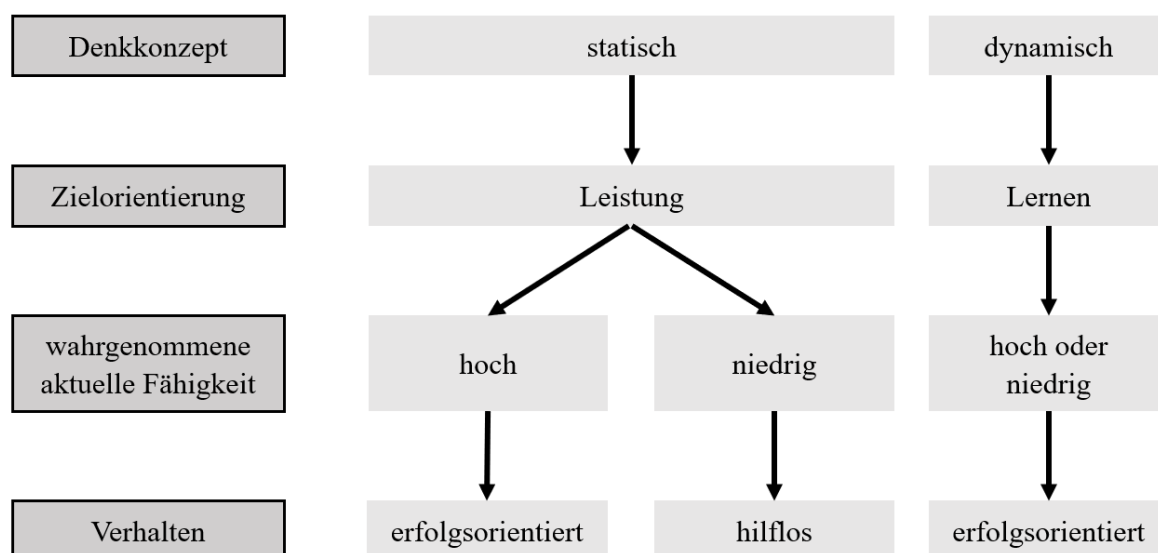


Abbildung 2. Darstellung des Zusammenhangs zwischen Denkkonzept und Verhalten nach Dweck (1986).

Kinder, die ein statisches Denkkonzept haben, streben positive Leistungsbeurteilungen an und möchten negative vermeiden, Kinder mit einem dynamischen Denkkonzept möchten ihre Fähigkeiten erweitern und dazulernen. Welches Denkkonzept Kinder aufweisen, äußert sich z. B. darin, ob sie glauben, dass Intelligenz unveränderbar (statisch) oder formbar (dynamisch) ist. Bei hohen wahrgenommenen Fähigkeiten in einer aktuellen Aufgabe zeigen sich die gleichen Verhaltensmuster bei beiden Gruppen: Mit hoher Beharrlichkeit suchen alle Kinder Herausforderungen und möchten die Aufgabe meistern. Bei niedriger wahrgenommener Fähigkeit in der aktuellen Aufgabe zeigt sich allerdings ein Unterschied: Kinder mit statischem Denkkonzept zeigen hilfloses Verhalten, sie vermeiden die Herausforderung und sind wenig beharrlich (Dweck, 1986; Dweck & Leggett, 1988; Elliott & Dweck, 1988).

Das Modell von der Forschungsgruppe um Dweck wurde in zahlreichen Studien bestätigt und erweitert (Blackwell, Trzesniewski & Dweck, 2007; Cimpian et al., 2007). Das dynamische Denkkonzept führt zu stärkeren Lernzielen, mehr positiven Annahmen über Anstrengung und weniger fähigkeitsbasierten, hilflosen Attributionen (Blackwell et al., 2007). Besonders hervorzuheben ist die Langfristwirkung dieses dynamischen Ansatzes: Es zeigen sich bessere Leistungen (in dieser Studie im Fachbereich der Mathematik; Blackwell et al., 2007) über zwei Jahre hinweg und eine stärkere Leistungssteigerung im Zeitverlauf als bei einem statischem Denkkonzept. Darüber hinaus zeigen sich Unterschiede hinsichtlich des

Umgangs mit Misserfolgen: Während Schüler*innen mit statischem Denkkonzept Misserfolge auf zu geringe Fähigkeiten attribuieren, sehen die dynamisch Denkenden die Ursache in zu geringer Anstrengung. Letzteres führt in der Konsequenz wiederum dazu, dass Herausforderungen auch nach Niederlagen erneut aufgesucht und mit stärkerer Anstrengung angegangen werden (Hong, Chiu, Dweck, Lin & Wan, 1999).

2.2.2 Feedback und Denkkonzepte

Der Zusammenhang zwischen statischem und dynamischen Denkkonzept und Feedback lässt sich an zwei Ansatzpunkten ausmachen – zum einen an der Entwicklung der jeweiligen Denkkonzepte, zum anderen am Umgang mit Feedback in Abhängigkeit vom bei der Person etablierten Denkkonzept.

Feedback als Anstoß zur Entwicklung von Denkkonzepten. Neben dem direkten Lehren der statischen oder dynamischen Intelligenztheorie werden statische bzw. dynamische Denkkonzepte auch indirekt über Feedback vermittelt (Dweck, 2002). Lob der Intelligenz als stabile Personeigenschaft (Personfeedback) bewirkt einen Glauben an statische Denkkonzepte und führt zu einer Verringerung der Leistungsmotivation von Schüler*innen und mehr hilflosen Reaktionen, wohingegen anstrengungsbezogenes Lob (Prozesslob) ein dynamisches Denkkonzept fördert (Kamins & Dweck, 1999; Mueller & Dweck, 1998). Das jeweilige Denkkonzept von Kindern kann sogar in Abhängigkeit von der Art des Lobs durch ihre Eltern in Langzeitstudien vorhergesagt werden (für personbezogenes Lob vgl. Pomerantz und Kempner (2013), für prozessbezogenes Lob vgl. Gunderson et al. (2013)). Rattan, Good und Dweck (2012) konnten zeigen, dass Lehrkräfte mit statischem Denkkonzept geringere Erwartungen bei ihren Schüler*innen an deren eigene Leistung hervorrufen und sie demotivieren.

Cimpian und Markman (2011) zeigen, dass auch generalisierte Formulierungen beim Feedbackgeben ein statisches Denkkonzept fördern können. So führen Kinder die Aussage „Jungen sind gut in Mathematik“ eher auf eine stabile, internale, ohne Mühe erhaltene Eigenschaft zurück, als die Aussage „Johnny ist gut in Mathematik“. Kommt es jedoch nach einem generalisierten Feedback wie „Du bist eine gute Zeichnerin“ zu Kritik, so bezweifeln Kinder ihre Fähigkeiten und vermeiden die zuvor noch gelobte Fähigkeit (Cimpian et al., 2007). Exakte sprachliche Differenzierung und Spezifizierung könnte demnach möglicherweise zur Qualität von Feedback beitragen.

Umgang mit Feedback in Abhängigkeit vom Denkkonzept. Zum Umgang mit Feedback konnten Hong et al. (1999) zeigen, dass Personen mit dynamischem Denkkonzept negatives Feedback auf mangelnde Anstrengung statt auf mangelnde Fähigkeit attribuieren und Gegenmaßnahmen bei nicht zufriedenstellender Leistung eher als Personen mit statischem Denkkonzept einleiteten. Personen mit dynamischem Denkkonzept zeigen einen konstruktiveren Umgang mit Misserfolgen.

2.3 Theoretisches Modell dieser Arbeit und Hypothesen

Feedback kann belastend sein – sowohl für die feedbacknehmende als auch für die feedbackgebende Person. Negative Konsequenzen für die feedbacknehmende Person wurden in den Kapiteln 2.1 und 2.2 ausführlich beschrieben, für die feedbackgebende Person kann problematisch sein, dass sie das Gegenüber nicht verletzen möchte oder selbst Angst vor einer negativen Reaktion hat (Besieux, 2017; Donnelly & Kirk, 2015). Herauszufinden, wie Rückmeldungen gestaltet sein müssen, dass sie sowohl auf Seiten der feedbackgebenden als auch der feedbacknehmenden Person wirksam und belastungsarm sind, ist Ziel dieser Arbeit.

Herangehensweise in dieser Arbeit. Ziel von Feedback ist die Veränderung von Verhalten (McKimm, 2009). Damit dies gelingen kann, empfiehlt die Forschung zahlreiche sich unterscheidende Wirkfaktoren (vgl. Kapitel 2.1.2.2), denen bisherige Formulierungen häufig nicht gerecht werden (vgl. Kapitel 2.1.3). Diese Masterarbeit versucht hier einen Beitrag zur Vereinfachung zu leisten: Mittels einer anderen Formulierung von Feedback soll der*die Feedbackgebende auf sprachlicher Ebene in eine inhaltlich andere Ausrichtung gelenkt werden und so einige der Wirkfaktoren automatisch (als Konsequenz der Formulierung) erfüllen, insofern er*sie grammatikalisch korrekte Aussagen trifft. Statt dichotom mittels „Was findest du gut? Was schlecht?“ oder „Was hat dir gefallen? Was nicht?“ nach positivem vs. negativem Feedback zu fragen, soll der Feedbackgebende seine Rückmeldung anhand der Leitfragen „Was sollte beibehalten werden? Was verändert?“ entwickeln und damit ein bestärkendes bzw. veränderndes Feedback geben.

Ableitung der Hypothesen. Da die Formulierungsvariante als *beibehalten vs. verändern* grundlegend anders als das dichotome Konzept *gut vs. schlecht* ist und bislang wohl auch weniger verbreitet ist, wird davon ausgegangen, dass Teilnehmende mehr kognitive Kapazität für die Auseinandersetzung mit der Formulierung benötigen, die dann nicht mehr für

die Erarbeitung von Inhalten zur Verfügung steht. Es wird angenommen, dass weniger inhaltliche Aspekte rückgemeldet werden, wenn Feedback nach *beibehalten vs. verändern* formuliert wird (vgl. Abbildung 3, Hypothese H₂).

Hypothesen zum Feedbackgeben (untersucht in Studie 1)
H ₁ : In der BV-Bedingung zeigt sich eine positivere Einstellung zum Feedback als in der GS-Bedingung.
H ₂ : In der BV-Bedingung ist die Anzahl genannter Aspekte geringer als in der GS-Bedingung.
H ₃ : In der BV-Bedingung ist das Feedback konkreter als in der GS-Bedingung.
H ₄ : In der BV-Bedingung werden andere Feedbackinhalte genannt als in der GS-Bedingung und zwar: a) verhaltensbezogener b) weniger eigenschaftsbezogen

Abbildung 3. Hypothesen zum Feedbackgeben (Studie 1).

Es ist hervorzuheben, dass es sich bei bestärkendem bzw. veränderndem Feedback nicht um eine reine Umformulierung des dichotomen Formulierungskonzeptes handelt, sondern um eine andere Art, Feedback zu formulieren. Dies wird anhand dreier Konsequenzen der Formulierung deutlich.

Erstens ist bestärkendes bzw. veränderndes Feedback nicht zwangsweise mit einer Wertung verbunden. So kann im verändernden Feedback auch rückgemeldet werden, dass eine gute Verhaltensweise noch weiter ausgebaut werden sollte (z. B. „Du lächelst so aufmunternd – bitte mehr davon!“) oder eine prinzipiell gute Verhaltensweise in spezifischem Kontext unangebracht ist („Ich würde mir wünschen, dass du keine Witze mehr über mich in großen Gruppen machst. Solange wir unter uns sind, mach aber gerne weiterhin diese Witze.“). Damit bietet dieses Konzept mehr Raum für Spezifizierung sowie Differenzierung in der Formulierung von Rückmeldung (vgl. H₃ in Abbildung 3) und in der Folge auch mehr Ansatzpunkte für mögliche Verhaltensänderungen (vgl. H₃ in Abbildung 4).

Hypothesen zum Feedbacknehmen (untersucht in Studie 2)
H ₁ : In der BV-Bedingung zeigt sich eine positivere Einstellung zum Feedback als in der GS-Bedingung.
H ₂ : In der BV-Bedingung wird das Feedback als hilfreicher empfunden als in der GS-Bedingung.
H ₃ : In der BV-Bedingung führt Feedback zu konkreteren Vorsätzen als in der GS-Bedingung.
H ₄ : In der BV-Bedingung führt Feedback zu mehr Vorsätzen als in der GS-Bedingung.

Abbildung 4. Hypothesen zum Feedbacknehmen (Studie 2).

Zweitens sind grammatikalische Unterschiede konzeptinhärent. Im dichotomen Formulierungsschema kann man sowohl Eigenschaften („Ich finde gut, dass deine Stimme so angenehm ist.“), als auch Verhaltensweisen („Ich finde gut, dass du so laut sprichst.“) und Situationen („Ich finde den Raum gut.“) gut oder schlecht. Nach dem bestärkenden bzw. verändernden Formulierungskonzept können nur Verhaltensweisen rückgemeldet werden („Ich finde, der Dozent sollte beibehalten, dass er so laut spricht.“). Dies wird umso deutlicher, wenn man eine Rückmeldung formuliert wie „Die Dozentin sollte ihr angenehme Stimme beibehalten“ – rein aus Sprachgefühl erscheint diese Formulierung falsch; logisch argumentiert muss die Dozentin ihre Stimme beibehalten, da diese kein direkt veränderbares Verhalten darstellt. Feedbackgebende werden mittels der Formulierung dazu angehalten, verhaltens- und damit prozessbezogenes Feedback zu geben, was diverse positive Konsequenzen mit sich bringt (vgl. Kapitel 2.1.2) Diese Überlegung zum Feedbackgeben spiegelt sich in Hypothese H₄ (vgl. Abbildung 3) wider.

Drittens wird angenommen, dass bei der Formulierung als *beibehalten* vs. *verändern* die ausbleibende Wertung und starke Verhaltensorientierung des Feedbacks die Belastung der Feedbackgebenden sowie -nehmenden verringert und dadurch eine positivere Einstellung zum Prozess des Feedbacks erzielt wird – beim Feedbackgebenden gleichermaßen wie beim Feedbacknehmenden (Hypothese H₁ in beiden Studien, vgl. Abbildung 3 und Abbildung 4).

Da vor allem bei Personen mit geringer Selbstwertschätzung Feedback negative Konsequenzen mit sich bringt (vgl. Kapitel 2.1.2.1), könnte hier großer Spielraum für Verbesserung durch optimiertes Formulieren von Feedback sein. Bei geringer Selbstwertschätzung von Feedbacknehmenden ist zu vermuten, dass die Unterschiede zwischen den beiden Konzepten noch deutlicher zu finden sind, da vor allem auf diese Gruppe das dichotome Formulierungskonzept negative Auswirkungen zu haben scheint. Da jedoch zunächst die Wirksamkeit des bestärkenden bzw. verändernden Formulierungskonzeptes zu überprüfen ist, wird diese Vermutung nur als explorativer Forschungsgegenstand, nicht explizit als Hypothese aufgenommen.

3 Studie 1: Feedbackgeben

3.1 Methode

3.1.1 Untersuchungsdesign und -ablauf

In der Studie zu Feedbackgeben gaben die Teilnehmenden Feedback zu universitären Lehrveranstaltungen ab. Die unabhängige Variable war das Formulierungskonzept des Feedbacks mit zwei Ausprägungen, entweder *beibehalten* vs. *verändern* oder *gut* vs. *schlecht*. Das Formulierungskonzept wurde zwischen den Teilnehmenden variiert.

Als abhängige Variable wurde für die Hypothese H₁ die Einstellung zum Feedback als Differenz aus der Einstellung zum aktuellen Feedback in diesem konkreten Fall (abgefragt nach der Abgabe des Feedbacks) und der generellen Einstellung zum Feedback (abgefragt vor der Abgabe des Feedbacks) erhoben. Bei Hypothese H₂ war die abhängige Variable die Anzahl genannter Aspekte, operationalisiert als die Anzahl ausgefüllter Zeilen im Feedbackteil des Fragebogens. Für Hypothese H₃ wurde die Konkretheit der in freiem Format rückgemeldeten Aspekte untersucht. Die Konkretheit wurde nach dem Abstraktheitssuffix-Verfahren von Günther und Groeben (1978) kodiert. In Hypothese H₄ wurden als abhängige Variablen die Verhaltensbezogenheit (H_{4a}) sowie die Eigenschaftsbezogenheit (H_{4b}) der Feedbackinhalte ausgewertet. Die Kodierung orientierte sich am Aktionsquotienten von Busemann (1925). Um zu kontrollieren, ob die Teilnehmenden sich an die jeweils vorgegebene Formulierungsvariante hielten, wurde darüber hinaus die grammatikalische Korrektheit der Rückmeldungen kodiert. Details zur Kodierung werden in Kapitel 3.1.4 dargestellt.

In zwei Teilstichproben wurde die Untersuchung online durchgeführt, in zwei Teilstichproben in persona (siehe Kapitel 3.1.2), der Ablauf war allerdings identisch. Vorab erhielten die Teilnehmenden nur eine kurze Instruktion, die erläuterte, dass mittels des Fragebogens Feedback zur jeweiligen Veranstaltung erhoben wird und dieses Feedback zum einen der Veranstaltungsverbesserung dient, zum anderen aber auch im Rahmen einer Masterarbeit ausgewertet werden soll. Die Einladungsmail wird der Arbeit als Anhang A angehängt, die mündliche Instruktion, die bei den Teilstichproben in persona gegeben wurde, ist inhaltlich und vom Umfang damit identisch.

Nach der einleitenden Instruktion füllten die Teilnehmenden den jeweiligen Fragebogen aus (entweder online nach Klick auf den Link oder in persona nach dem Verteilen der Bögen). Die Fragebögen enthalten ausführliche Instruktionen und Abfragen des Einverständnisses

zur Teilnahme an der Studie. Alle Teilnehmenden erklärten der Teilnahme sowie der Auswertung der Daten im Rahmen dieser Masterarbeit einverstanden. Auf die Fragebögen wird in Kapitel 3.1.3 näher eingegangen.

Alle Teilnehmenden – sowohl die Online- als auch die in persona Teilnehmenden – wurden nach der Erhebung per E-Mail über das Untersuchungsdesign aufgeklärt. Kontaktdaten der Untersuchungsleitung wurden auf dem Fragebogen und zusätzlich in der Aufklärungsnachricht angegeben.

Das abgegebene Feedback der Teilnehmenden wurde in aggregierter Form den jeweiligen Dozierenden rückgemeldet.

3.1.2 Stichprobe

Insgesamt 78 Studierende nahmen freiwillig an der Erhebung teil (geworben wurden sie mittels Anhang A), davon 44 in der Bedingung, die Feedback nach der Formulierung *beibehalten vs. verändern* gegeben hat (BV-Bedingung) und 34 in der Bedingung, die Feedback nach der Formulierung *gut vs. schlecht* gegeben hat (GS-Bedingung). Die Stichprobe war bezogen auf das Geschlecht ausgeglichen (51.3% weiblich) und alle Teilnehmenden waren Studierende, wobei wirtschaftliche bzw. informationstechnologische¹ Studiengänge (mit 67.9%) und Lehramt (mit 17.9%) die am häufigsten vertretenen Studiengänge waren.

Von den 78 erhobenen Fragebögen erfolgte mit 55 Rückmeldungen der Großteil im Rahmen einer Wirtschaftsinformatik-Übung. Neun Fragebögen stammten aus einem Psychologie-Seminar. Diese beiden Erhebungen wurden in persona durchgeführt, das heißt insgesamt wurden 82% der Rückmeldungen mit Papier-Bleistift-Methode vor Ort in der Veranstaltung erhoben. Die übrigen 18% der Fragebögen, die online erhoben wurden, stammten aus einem Examensvorbereitungskurs der Literaturwissenschaft (9 Fragebögen) sowie aus einem Lehramtskurs für Sachunterricht (5 Fragebögen).

Zusätzlich zur Möglichkeit, durch Feedback zur Verbesserung der Veranstaltung beizutragen, wurde den Teilnehmenden kein Anreiz geboten.

¹ Wirtschaftlich bzw. informationstechnologisch wird in diesem Fall zusammengefasst, da viele der Studierenden einen Schnittstellen-Studiengang besuchen, z. B. Wirtschaftsinformatik oder Betriebswirtschaftslehre mit Schwerpunkten im IT-Bereich.

Die Zuweisung zu den verschiedenen Experimentalbedingungen (BV-Bedingung oder GS-Bedingung) erfolgte online randomisiert, in persona nach der Sitzordnung. Eine randomisierte und bezüglich der Anzahl gleichverteilte Zuordnung war nicht möglich, da direkt nebeneinandersitzende Teilnehmende den gleichen Fragebogen erhalten sollten (um zu verhindern, dass sie zufällig beim Sitznachbarn die andere Formulierung entdecken und ihnen dann die experimentelle Manipulation bewusst wird). Stattdessen wurde etwa in der Mitte des Raumes an einer Reihe freier Sitzplätze geteilt und eine Seite der BV-Bedingung, die andere der GS-Bedingung zugewiesen.

Für die Hypothesen H₃ zur Konkretheit und H₄ zur Verhaltens- und Eigenschaftsbezogenheit des Feedbacks mussten zwei Datensätze ausgeschlossen werden, da keinerlei Aspekte rückgemeldet wurden, deren Konkretheit, Verhaltens- oder Eigenschaftsbezogenheit hätte bewertet werden können. Für die übrigen beiden Hypothesen wurden diese Datensätze nicht ausgeschlossen, da sowohl die Einstellung der Teilnehmenden zum Feedback erhoben werden konnte als auch die Anzahl der rückgemeldeten Aspekte (nämlich keine) gewertet werden konnte.

In der BV-Bedingung befanden sich 44 Teilnehmende, in der GS-Bedingung 34 [32]² Teilnehmende. Die beiden Bedingungen unterschieden sich weder nennenswert hinsichtlich der Geschlechterzusammensetzung (BV-Bedingung: 50.0% weiblich, GS-Bedingung: 52.9% weiblich), noch der Studienrichtungen (BV-Bedingung: 70.5% wirtschaftlich bzw. informationstechnologischer Studiengang, 18.2% Lehramt; GS-Bedingung: 64.7% [68.8%] wirtschaftlich bzw. informationstechnologischer Studiengang, 17.6% [15.6%] Lehramt) oder der Erhebungsformen (BV-Bedingung: 18.2% online, GS-Bedingung: 17.6% [15.6%] online).

Die Teilnehmenden zeigten durchschnittlich in der BV-Bedingung ($M_{BV} = 4.98$, $SD_{BV} = 1.56$, $n_{BV} = 44$) eine etwas positivere generelle Einstellung zu Feedback als in der GS-Bedingung ($M_{GS} = 4.47$, $SD_{GS} = 1.48$, $n_{GS} = 34$). Dieser Unterschied war nicht signifikant, $t(72.81) = -1.46$, $p = .148$, $d = -0.34$. Auch wenn dementsprechend keine Differenzbildung zur Berechnung der abhängigen Variable in Hypothese H₁ notwendig wäre (weil kein statistisch signifikanter Unterschied in den Bedingungen besteht in der generellen Einstellung gegenüber Feedback), wird dennoch wie in Kapitel 3.1.1 beschrieben darauf zurückgegriffen, um bei der recht kleinen Stichprobe auch minimale Schwankungen zu berücksichtigen.

² In eckigen Klammern stehen die analogen Kennwerte für die in den Hypothesen H₃ und H₄ verwendete Stichprobe, insofern sich diese vom vollständigen Datensatz unterscheiden.

Die Prüfung der grammatikalischen Korrektheit der Rückmeldungen ergab einen Anteil teilweise korrekter (bzw. falscher) Rückmeldungen von 18.8% (bzw. 3.1%) in der BV-Bedingung und 13.6% (bzw. 6.8%) in der GS-Bedingung. Dies stellte keinen signifikanten Unterschied zwischen den beiden Bedingungen dar, $\chi^2(2) = 0.82, p = .663$.

3.1.3 Fragebögen

Die Papier-Bleistift-Fragebögen waren fünfseitig, die Online-Erhebung wies identischen Wortlaut und gleiche Reihenfolge auf. In Anhang B ist der vollständige Fragebogen für die BV-Bedingung zu finden, in Anhang C die sich davon unterscheidenden Seiten für die GS-Bedingung. Nach den Einverständniserklärungen zur Teilnahme an der Studie, zur Verarbeitung der Daten im Rahmen der Masterarbeit und der Veröffentlichung der Daten in einem online verfügbaren wissenschaftlichen Datenarchiv folgten Items zu soziodemografischen Angaben. Anschließend wurde auf einer siebenstufigen Ratingskala (1 = *überhaupt nicht*, 7 = *voll und ganz*) die Zustimmung zur Aussage „Ich gebe generell gerne Feedback“ erhoben.

Es folgte eine Seite, auf der je nach Bedingung 15 Mal der Satzanfang „Ich finde, es sollte beibehalten werden, dass“ bzw. „Ich finde gut, dass“ notiert war und eine weitere Seite mit den Satzanfängen „Ich finde, es sollte verändert werden, dass“ bzw. „Ich finde schlecht, dass“.

Zum Abschluss wurde wiederum auf einer siebenstufigen Ratingskala (1 = *überhaupt nicht*, 7 = *voll und ganz*) die Zustimmung zur Aussage „Ich habe dieses Mal gerne Feedback gegeben“ erhoben. In einem Freitextfeld darunter erhielten die Teilnehmenden die Möglichkeit, Gründe für ihre Angabe zu erläutern. Dies diente als zusätzliche Informationsquelle für den Fall, dass Ausreißer aufgetreten wären.

3.1.4 Kodierung

Für die Hypothese H₃ zur Konkretetheit der rückgemeldeten Aspekte wurden die Rückmeldungen nach dem Abstraktheitssuffix-Verfahren von Günther und Groeben (1978) für deutsche Begriffe sowie nach dem analogen englischsprachigen Verfahren von Gillie (1957) für englische Begriffe ausgewertet. Substantive mit den Endungen *-heit*, *-ie*, *-ik*, *-ion*, *-ismus*, *-tät*, *-keit*, *-nz*, *-tur* und *-ung* werden als Indikatoren für die Abstraktheit eines Textes

gewertet. Die Split-half-Reliabilität des Abstraktheitssuffix-Verfahrens liegt in der Untersuchung von Günther und Groeben (1978) bei .94. Die Korrelation mit Lesereinschätzungen der Abstraktheit von Texten als Maß der Validität liegt bei .75 (Günther & Groeben, 1978). Allerdings ist zu berücksichtigen, dass Günther und Groeben (1978) von einer Mindestanzahl von 400 Substantiven pro Text ausgehen – in dieser Studie lag die Anzahl der Substantive pro Teilnehmer*in zwischen 1 und 28 ($M = 9.87$, $SD = 4.81$). Das Verfahren wurde dennoch herangezogen, da in dieser ersten Annäherung an die Thematik auf ein bewährtes Verfahren zurückgegriffen werden sollte.

Die über dieses Verfahren bestimmte Anzahl an Abstraktheitssubstantiven wurde dann von der Gesamtzahl der verwendeten Substantive subtrahiert. Dies ergab die Anzahl an konkreten Substantiven. Relativiert an der Gesamtzahl der verwendeten Substantive pro Fragebogen ergibt sich somit die Konkretheit der Rückmeldung eines*r Teilnehmenden.

Für die Hypothese H_{4a} zur Verhaltensbezogenheit von Aussagen wurde die Anzahl aktionaler Aussagen an der Anzahl aller rückgemeldeten Aspekte (also benutzter Zeilen) relativiert. Dabei sind aktionale Aussagen nach Busemann (1925) Aussagen mit finitem Verb und im Infinitiv mit (*um*) *zu* bzw. Infinitive mit den Hilfsverben *können*, *wollen*, *mögen*, *sollen* und *dürfen*, insofern sie nicht als qualitative Aussagen gewertet werden müssen.

Für die Hypothese H_{4b} zur Eigenschaftsbezogenheit von Aussagen wurde wiederum die Anzahl ebendieser qualitativen Aussagen an der Anzahl aller rückgemeldeten Aspekte (also benutzter Zeilen) relativiert. Qualitative Aussagen sind nach Busemann (1925) alle Aussagen mit dem Hilfszeitwort *sein*, alle Appositionen und attributiven Adjektive, Ordinalzahlen, Partizipien und Substantive. Hinzugefügt wurden in dieser Arbeit im Hinblick auf die interessierende Variable der Eigenschaftsbezogenheit alle adverbialen Adjektive.

Darüber hinaus wurde die grammatikalische Korrektheit des Rückmeldungsstils je Teilnehmer*in kodiert, die als gegeben galt, wenn die Angabe der*des Teilnehmenden eine grammatikalisch korrekte Ergänzung des jeweiligen Satzanfangs (z. B. „Ich finde gut, dass“) war. Es war neben der Einschätzung als *korrekt* auch die Einschätzung *teilweise korrekt* oder *falsch* möglich. *Teilweise korrekt* wurde bei denjenigen Teilnehmenden vermerkt, die zum Teil grammatikalisch korrekte und zum Teil falsche Aussagen trafen.

Die Interraterreliabilität, bestimmt über Cohens κ , für dieses Kodierverfahren lag in dieser Arbeit nach Doppelkodierung von 10% des Materials bei $\kappa = .88$, was nach Landis und Koch

(1977) einer nahezu vollkommenen bzw. nach Greve, Wentura, Gräser und Schmitz (1997) einer ausgezeichneten Übereinstimmung entspricht.

Die Kodierregeln sind im Detail in Anhang D zu finden.

3.2 Ergebnisse

3.2.1 Einstellung zum Feedback

In Hypothese H₁ wurde angenommen, dass sich in der BV-Bedingung eine positivere Einstellung zum Feedback zeigt. Wie in den Kapiteln 3.1.1 und 3.1.2 erläutert, wurde hierzu die Differenz aus der Einstellung zum aktuellen Feedback und der Einstellung zu Feedback generell herangezogen.

Voraussetzungsprüfung für t-Test (nach Bortz & Schuster, 2016). Die Normalverteilungsannahme für die Einstellungs­differenz gegenüber dem Feedback ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.91$ und $p_{BV} = .002$ sowie $W_{GS} = 0.85$ und $p_{GS} < .001$, da der t-Test jedoch gegenüber dieser Annahmeverletzung recht robust ist (Boneau, 1960) und darüber hinaus die Bedingungen jeweils $n > 30$ aufweisen, kann er trotzdem herangezogen werden. Die zufällige Zuweisung zu den Untersuchungsbedingungen ist insofern gegeben, dass keine inhaltlich bedeutsame Zuweisung erfolgte. Die Stichproben sind unabhängig voneinander. Da die Varianzhomogenität nach dem Levene-Test nicht verletzt ist, $F(1, 76) = 0.29$, $p = .590$, ist es unproblematisch, dass sich die Anzahl der Teilnehmenden je Bedingung unterschied ($n_{BV} = 44$, $n_{GS} = 34$). Ein t-Test ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel gaben die Teilnehmenden im aktuellen Fall – bereinigt um ihre generelle Einstellung zum Feedback – in der BV-Bedingung ($M_{BV} = -0.52$, $SD_{BV} = 1.49$, $n_{BV} = 44$) weniger gerne Feedback als in der GS-Bedingung ($M_{GS} = -0.35$, $SD_{GS} = 1.98$, $n_{GS} = 34$), siehe Tabelle 1. Dieser Mittelwertsunterschied widerspricht der gerichteten Hypothese und wird daher unter dem Punkt *Explorative Prüfungen* näher beleuchtet. Die Hypothese H₁, dass sich in der BV-Bedingung eine positivere Einstellung zum Feedback zeigt, muss demnach verworfen werden.

Tabelle 1. Mittelwert und Standardabweichung der Einstellung zum Feedbackgeben generell, im aktuellen Fall und als Differenz

Variable	Untersuchungsbedingung			
	BV-Bedingung ^d		GS-Bedingung ^e	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Generell ^a	4.98	1.56	4.47	1.48
Aktuell ^b	4.45	2.19	4.12	1.79
Differenz ^c	-0.52	1.49	-0.35	1.98

Anmerkungen. Die Zustimmung zum jeweiligen Item wurde auf einer Ratingskala erhoben (1 = überhaupt nicht, 7 = voll und ganz), $n_{BV} = 44$ und $n_{GS} = 34$.

^a Das Item lautete „Ich gebe generell gerne Feedback“. ^b Das Item lautete „Ich habe dieses Mal gerne Feedback gegeben“. ^c Berechnet wurde je Teilnehmer*in die Einstellung zum aktuellen Feedback abzüglich der Einstellung zum Feedback generell. ^d BV-Bedingung steht für die Bedingung, in der die Teilnehmenden Feedback in der Variante *beibehalten vs. verändern* geben. ^e GS-Gruppe steht für die Gruppe, in der die Teilnehmenden Feedback in der Variante *gut vs. schlecht* geben.

Explorative Prüfungen. Der unter *Ergebnisse der Hypothesenprüfung* beschriebene Mittelwertsunterschied hinsichtlich der Einstellung zum Feedback zwischen der BV-Bedingung und der GS-Bedingung wird in der explorativ durchgeführten ungerichteten Testung nicht signifikant, $t(59.29) = 0.42, p = .678, d = 0.11$.

Bei Betrachtung der Mittelwerte in Tabelle 1 ist auffällig, dass in beiden Bedingungen die Differenz zwischen der Einstellung zum aktuellen Feedback und der generellen Einstellung zu Feedback negativ ist, d.h. dass durchschnittlich im aktuellen Fall deutlich weniger gerne Feedback gegeben wurde als generell. Dieser Unterschied ist in der BV-Bedingung auf einem nach Bonferroni-Korrektur adjustiertem Alphafehlerniveau von $\alpha = .05/2 = .025$ signifikant, $t(43) = -2.33, p = .024, d = -0.71$, in der GS-Bedingung hingegen nicht, $t(33) = -1.04, p = .307, d = -0.36$. Die Alphafehler-Korrektur wurde in dieser Testung vorgenommen, da die gleiche abhängige Variable in zwei Tests (einmal innerhalb der BV- und einmal innerhalb der GS-Bedingung) inferenzstatistisch überprüft wurde.

3.2.2 Anzahl der genannten Aspekte

In Hypothese H₂ wurde angenommen, dass in der BV-Bedingung weniger Aspekte rückgemeldet werden als in der GS-Bedingung, da dieses Konzept neu für die Teilnehmenden ist. Wie in Kapitel 3.1.1 erläutert, wurde hierzu die Anzahl ausgefüllter Zeilen im Feedbackteil des Fragebogens herangezogen.

Voraussetzungsprüfung für t-Test (nach Bortz & Schuster, 2016). Die Normalverteilungsannahme für die Anzahl der genannten Aspekte ist in keiner der Bedingungen nach dem Shapiro-Wilk-Test verletzt, $W_{BV} = 0.96$ und $p_{BV} = .144$ sowie $W_{GS} = 0.96$ und $p_{GS} = .308$. Die zufällige Zuweisung zu den Untersuchungsbedingungen ist insofern gegeben, dass keine inhaltlich bedeutsame Zuweisung erfolgte. Die Stichproben sind unabhängig voneinander. Die Verletzung der Varianzhomogenität nach dem Levene-Test, $F(1, 76) = 7.54$, $p = .008$ und die sich unterscheidende Anzahl der Teilnehmenden je Bedingung ($n_{BV} = 44$, $n_{GS} = 34$) fielen wegen der Verwendung des gegen diese Verletzung robusten Zweistichproben-t-Test nach Welch nicht negativ ins Gewicht (Ruxton, 2006; Welch, 1938). Ein Zweistichproben-t-Test nach Welch ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel wurden etwas weniger Aspekte in der BV-Bedingung ($M_{BV} = 4.86$, $SD_{BV} = 2.03$, $n_{BV} = 44$) genannt als in der GS-Bedingung ($M_{GS} = 5.71$, $SD_{GS} = 3.41$, $n_{GS} = 34$). Dieser Mittelwertsunterschied ist nicht signifikant $t(50.66) = 1.28$, $p = .104$, $d = 0.36$. Eine Sensitivitätsanalyse ergab, dass unter den gegebenen Parametern und Stichprobengrößen der Effekt $d \geq .58$ hätte sein müssen, um ihn – insofern er in der Population tatsächlich existiert – auch vorfinden zu können. Die Teststärkebestimmung a posteriori ergab eine Teststärke von 46.7%.

Die Hypothese H_2 , dass in der BV-Bedingung weniger Aspekte rückgemeldet werden, muss verworfen werden.

3.2.3 Konkretheit des Feedbacks

In Hypothese H_3 wurde angenommen, dass in der BV-Bedingung das Feedback konkreter ist. Wie in den Kapiteln 3.1.1 und 3.1.4 erläutert, wurde die Anzahl an Abstraktheitssubstantiven (nach Günther & Groeben, 1978) von der Anzahl aller Substantive subtrahiert und dieser Wert der konkreten Substantive an der Anzahl aller Substantive relativiert. Dies ergab die Konkretheitsrate je Teilnehmer*in.

Voraussetzungsprüfung für t-Test (nach Bortz & Schuster, 2016). Die Normalverteilungsannahme für die Konkretheitsrate ist in der GS-Bedingung ($W_{GS} = 0.93$, $p_{GS} = .033$) nach dem Shapiro-Wilk-Test signifikant verletzt, in der BV-Bedingung nicht ($W_{BV} = 0.95$, $p_{BV} = .083$), da der t-Test jedoch gegenüber dieser Annahmeverletzung recht robust ist

(Boneau, 1960) und darüber hinaus die Bedingungen jeweils $n > 30$ aufweisen, kann er trotzdem herangezogen werden. Die zufällige Zuweisung zu den Untersuchungsbedingungen ist insofern gegeben, dass keine inhaltlich bedeutsame Zuweisung erfolgte. Die Stichproben sind unabhängig voneinander. Da die Varianzhomogenität nach dem Levene-Test nicht verletzt ist, $F(1, 74) = 0.08, p = .776$, ist es unproblematisch, dass sich die Anzahl der Teilnehmenden je Bedingung unterschied ($n_{BV} = 44, n_{GS} = 32$). Ein t-Test ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel war die Konkrettheitsrate in der BV-Bedingung ($M_{BV} = 0.65, SD_{BV} = 0.20, n_{BV} = 44$) etwas niedriger als in der GS-Bedingung ($M_{GS} = 0.73, SD_{GS} = 0.20, n_{GS} = 32$). Dieser Mittelwertsunterschied widerspricht der gerichteten Hypothese und wird daher unter dem Punkt *Explorative Prüfungen* näher beleuchtet. Die Hypothese H_3 , dass in der BV-Bedingung das Feedback konkreter ist, muss demnach verworfen werden.

Explorative Prüfungen. Der unter *Ergebnisse der Hypothesenprüfung* beschriebene Mittelwertsunterschied hinsichtlich der Konkrettheitsrate in der BV-Bedingung und der GS-Bedingung wird in der explorativ durchgeführten ungerichteten Testung nicht signifikant, $t(67.35) = 1.68, p = .098, d = 0.41$.

Die Prüfung der Hypothese H_3 auf Basis eines reduzierten Datensatzes (nach Ausschluss aller Fälle mit nur teilweise korrekter oder falscher Grammatik) zeigt ebenso keinen signifikanten Unterschied zwischen der BV-Bedingung ($M_{BV} = 0.67, SD_{BV} = 0.20, n_{BV} = 35$) und der GS-Bedingung ($M_{GS} = 0.74, SD_{GS} = 0.21, n_{GS} = 24$), $t(48.29) = 1.36, p = .181, d = 0.39$, und weist auffällig ähnliche deskriptive und inferenzstatistische Werte auf wie die Auswertung im vollständigen Datensatz.

3.2.4 Verhaltensbezogenheit der Feedbackinhalte

In Hypothese H_{4a} wurde angenommen, dass in der BV-Bedingung stärker verhaltensbezogenes Feedback gegeben wird. Wie in Kapitel 3.1.4 erläutert, wurde hierzu die Rate aktionaler Aussagen herangezogen.

Voraussetzungsprüfung (nach Bortz & Schuster, 2016). Die Normalverteilungsannahme für die Rate aktionaler Aussagen ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.89$ und $p_{BV} < .001$ sowie $W_{GS} = 0.91$ und $p_{GS} = .011$, da der t-Test jedoch gegenüber dieser Annahmeverletzung recht robust ist (Boneau, 1960) und

darüber hinaus die Bedingungen jeweils $n > 30$ aufweisen, kann er trotzdem herangezogen werden. Die zufällige Zuweisung zu den Untersuchungsbedingungen ist insofern gegeben, dass keine inhaltlich bedeutsame Zuweisung erfolgte. Die Stichproben sind unabhängig voneinander. Da die Varianzhomogenität nach dem Levene-Test nicht verletzt ist, $F(1, 74) = 0.72, p = .399$, ist es unproblematisch, dass sich die Anzahl der Teilnehmenden je Bedingung unterschied ($n_{BV} = 44, n_{GS} = 32$). Ein t-Test ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel war die Verhaltensbezogenheit des Feedbacks in der BV-Bedingung ($M_{BV} = 0.84, SD_{BV} = 0.36, n_{BV} = 44$) etwa genauso groß wie in der GS-Bedingung ($M_{GS} = 0.85, SD_{GS} = 0.28, n_{GS} = 32$). Diese Ähnlichkeit widerspricht der gerichteten Hypothese und wird daher unter dem Punkt *Explorative Prüfungen* näher beleuchtet. Die Hypothese H_{4a} , dass in der BV-Bedingung stärker verhaltensbezogen Feedback gegeben wird, muss demnach verworfen werden.

Explorative Prüfungen. Der unter *Ergebnisse der Hypothesenprüfung* beschriebene Mittelwertsunterschied hinsichtlich der Verhaltensbezogenheit des Feedbacks wird in der explorativ durchgeführten ungerichteten Testung nicht signifikant, $t(73.50) = 0.20, p = .845, d = 0.05$.

Die Prüfung der Hypothese H_{4a} auf Basis eines reduzierten Datensatzes (nach Ausschluss aller Fälle mit nur teilweise korrekter oder falscher Grammatik) zeigt ebenso keinen signifikanten Unterschied zwischen der BV-Bedingung ($M_{BV} = 0.91, SD_{BV} = 0.3, n_{BV} = 35$) und der GS-Bedingung ($M_{GS} = 0.92, SD_{GS} = 0.23, n_{GS} = 24$), $t(56.04) = 0.04, p = .969, d = 0.01$, und weist auffällig ähnliche deskriptive und inferenzstatistische Werte auf wie die Auswertung im vollständigen Datensatz.

3.2.5 Eigenschaftsbezogenheit der Feedbackinhalte

In Hypothese H_{4b} wurde angenommen, dass in der BV-Gruppe weniger eigenschaftsbezogen Feedback gegeben wird. Wie in Kapitel 3.1.4 erläutert, wurde hierzu die Rate qualitativer Aussagen herangezogen.

Voraussetzungsprüfung für t-Test (nach Bortz & Schuster, 2016). Die Normalverteilungsannahme für die Rate qualitativer Aussagen ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.88$ und $p_{BV} < .001$ sowie $W_{GS} = 0.91$ und $p_{GS} = .009$, da der t-Test jedoch gegenüber dieser Annahmeverletzung recht robust ist

(Boneau, 1960) und darüber hinaus die Bedingungen jeweils $n > 30$ aufweisen, kann er trotzdem herangezogen werden. Die zufällige Zuweisung zu den Untersuchungsbedingungen ist insofern gegeben, dass keine inhaltlich bedeutsame Zuweisung erfolgte. Die Stichproben sind unabhängig voneinander. Da die Varianzhomogenität nach dem Levene-Test nicht verletzt ist, $F(1, 74) = 0.07, p = .797$, ist es unproblematisch, dass sich die Anzahl der Teilnehmenden je Bedingung unterschied ($n_{BV} = 44, n_{GS} = 32$). Ein t-Test ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel war die Eigenschaftsbezogenheit des Feedbacks in der BV-Bedingung ($M_{BV} = 0.69, SD_{BV} = 0.47, n_{BV} = 44$) geringer als in der GS-Bedingung ($M_{GS} = 0.92, SD_{GS} = 0.48, n_{GS} = 32$). Dieser Mittelwertsunterschied ist signifikant $t(65.53) = 2.10, p = .020, d = 0.52$. Die Hypothese H_{4b} , dass in der BV-Bedingung weniger eigenschaftsbezogen Feedback gegeben wird, kann beibehalten werden.

Explorative Prüfungen. Die Prüfung der Hypothese H_{4b} auf Basis eines reduzierten Datensatzes (nach Ausschluss aller Fälle mit nur teilweise korrekter oder falscher Grammatik) zeigt ebenso einen signifikanten Unterschied zwischen der BV-Bedingung ($M_{BV} = 0.67, SD_{BV} = 0.46, n_{BV} = 35$) und der GS-Bedingung ($M_{GS} = 0.97, SD_{GS} = 0.53, n_{GS} = 24$), $t(58) = 12.11, p = .022, d = 3.18$, und weist auffällig ähnliche deskriptive und noch deutlichere inferenzstatistische Werte auf wie die Auswertung im vollständigen Datensatz.

3.3 Diskussion

In Studie 1 wurde untersucht, ob die Formulierung von Feedback *beibehalten vs. verändern* annehmbarere, spezifischere und prozessbezogenere Rückmeldungen seitens der Feedbackgebenden generiert als die Formulierung *gut vs. schlecht*. Lediglich letzteres lässt sich zum Teil bestätigen, da die Formulierung *beibehalten vs. verändern* in dieser Studie zu signifikant weniger eigenschaftsbezogenem Feedback führt.

Im Folgenden sollen zunächst die Ergebnisse der Hypothesenprüfung für diese Studie einzeln diskutiert, in den aktuellen Forschungsstand eingeordnet und Ansatzpunkte für zukünftige Forschung dargestellt werden. In Kapitel 5 werden anschließend die Ergebnisse beider Studien zusammengeführt und ein Gesamtausblick gegeben.

Die in Hypothese H_1 angenommene positivere Einstellung zum Feedback in der BV-Bedingung ließ sich nicht bestätigen, es zeigte sich kein signifikanter Unterschied zwischen den beiden Bedingungen. Eine auf diese Art und Weise operationalisierte geringere Belastung

des Feedbackgebenden (Besieux, 2017; Donnelly & Kirk, 2015) durch das Formulierungskonzept *beibehalten vs. verändern* im Gegensatz zu *gut vs. schlecht* lässt sich anhand der vorliegenden Ergebnisse nicht nachweisen. Grund dafür könnte die direkte Erfassung der Einstellung gegenüber dem Feedback mittels 1-Item-Abfrage („Ich habe dieses Mal gerne Feedback gegeben“) sein. Viele Teilnehmende kreuzten dieselbe Angabe an wie im Kontrollitem „Ich habe dieses Mal gerne Feedback gegeben“ und begründeten diese Auswahl auch explizit mit dem Kommentar „Ich gebe immer gerne Feedback“. Möglicherweise wäre hier eine andere Erfassungsmethode zielführender, entweder über implizite Methoden zur Einstellungsmessung, über den Affekt oder über eine explizite Erfassung mit „Im Vergleich zu sonst habe ich dieses Mal ... Feedback gegeben“ und einer Ratingskala zwischen gar nicht gerne über genauso gerne bis sehr gerne.

Ein signifikanter Unterschied zeigt sich zwischen der Einstellung zum Feedbackgeben generell und im aktuellen Fall: Die Teilnehmenden in der BV-Bedingung gaben aktuell signifikant weniger gerne Feedback als generell. Grund dafür könnte das ungewohnte Feedbackformat sein: Es ist anstrengender und aufwändiger, nach einem unbekanntem neuen Format Feedback zu formulieren – was jedoch nicht bedeutet, dass dies nicht möglich wäre (vgl. Sluijsmans et al., 2002).

Auffällig ist die Standardabweichung bei der aktuellen Einstellung zum Feedback in der BV-Bedingung: Verglichen mit den Standardabweichungen der generellen Einstellung zum Feedback ist sie höher (siehe Tabelle 1). Dies bietet einen Ansatzpunkt für Folgestudien, in denen untersucht werden könnte, welche Personengruppen besonders gerne Feedback nach *beibehalten vs. verändern* geben und welche nicht.

In Hypothese H₂ wurde geprüft, ob im Rahmen des neuen Formulierungskonzeptes die Anzahl an Rückmeldungen geringer ist. Der hypothesenkonforme Mittelwertsunterschied wurde jedoch nicht signifikant. Die geringe Teststärke – die im Übrigen für alle geprüften Hypothesen in Studie 1 zu finden ist – ist in diesem Fall besonders kritisch zu betrachten. Unter gegebenen Parametern und Stichprobengrößen lag die Wahrscheinlichkeit eines signifikanten Testergebnisses bei tatsächlicher Gültigkeit der postulierten Hypothese bei nur 46.7%. Für zukünftige Studien sollte eine den zu erwartenden Effekten angemessenere Stichprobengröße untersucht werden.

Die Prüfung der Hypothese H₃ zeigt, dass in der Konkretheit des Feedbacks kein signifikanter Unterschied zwischen den beiden Bedingungen besteht. Die Annahme, dass das Formulierungskonzept nach *beibehalten* vs. *verändern* mehr Raum für Spezifizierung sowie Differenzierung bietet, kann nicht aufrechterhalten werden. Kritisch zu bewerten ist in diesem Zusammenhang die Verwendung des Abstraktheitssuffix-Verfahrens (Günther & Groben, 1978) zur Kodierung der Konkretheit der rückgemeldeten Aspekte. Dieses Verfahren ist ausgelegt für eine Mindestanzahl von 400 Substantiven pro Text, wobei die in dieser Studie zu bewertenden Textstücke im Mittel etwa zehn Substantive enthielten (siehe Kapitel 3.1.4). Dieses Verfahren wurde nach intensiver Prüfung von Alternativen dennoch herangezogen, da es ein standardisiertes, reliables und valides Maß (bei Berücksichtigung der Anwendungsbedingungen, die hier nur bedingt gegeben waren) darstellt. In zukünftiger Forschung zur Thematik ist dies zu berücksichtigen und ein zutreffenderes Maß auszuwählen bzw. zu entwickeln.

Die in Hypothese H₄ angenommenen Unterschiede bezüglich der Inhalte des Feedbacks wurden zum Teil bestätigt. So zeigte sich zwar bzgl. der Verhaltensbezogenheit (in H_{4a}) kein signifikantes Ergebnis, aber bzgl. der geringeren Eigenschaftsbezogenheit (in H_{4b}) schon: In der BV-Bedingung wurden wie erwartet signifikant weniger eigenschaftsbezogene Inhalte rückgemeldet. Dieses Ergebnis zeigt, dass mittels einer anderen Formulierung in der Tat andere Inhalte zurückgemeldet werden und damit das Formulierungskonzept *beibehalten* vs. *verändern* inhaltlich wirksameres – nämlich weniger personbezogenes (Donnelly & Kirk, 2015; Evans, 2013; Gibbs & Simpson, 2004; Pomerantz & Kempner, 2013) – Feedback generiert.

Hypothesenübergreifend sollen an dieser Stelle zwei grundsätzliche Limitationen der Studie dargelegt werden: Trotz a priori festgestellter Stichprobengröße konnte keine angemessen große Stichprobe erzielt werden, sondern nur 60.9% [59.4%] der a priori berechneten Teilnehmendenzahl. Dies ist im Wesentlichen durch die zeitlichen und finanziellen Rahmenbedingungen einer Masterarbeit bedingt. Für zukünftige Untersuchungen in diesem Forschungsbereich sollte mit den zu erwartenden Effekten angemessen großen Stichproben gearbeitet werden.

Darüber hinaus ist festzustellen, dass der Anteil grammatikalisch – teilweise oder vollständig – falscher Rückmeldenden mit etwa $\frac{1}{4}$ aller Teilnehmenden recht hoch war. Nachdem zu diskutieren ist, ob in diesen Fällen die Manipulation der Teilnehmenden, Feedback entweder

nach der einen oder nach der anderen Variante zu geben, als geglückt gewertet werden kann, wurden neben den Prüfungen mit dem vollständigen Datensatz zusätzlich explorative Berechnungen mit einem reduzierten Datensatz durchgeführt. Die explorativen Prüfungen der Hypothesen H_3 und H_4 mit Ausschluss der grammatikalisch nicht korrekten Fälle zeigten jedoch keine nennenswerten Unterschiede zu den Tests, in denen alle Fälle einbezogen wurden.

4 Studie 2: Feedbacknehmen

4.1 Methode

4.1.1 Untersuchungsdesign

In der Studie zu Feedbacknehmen erhielten die Teilnehmenden Feedback von geschulten Beobachtenden zu ihrem Verhalten in Kooperationsaufgaben. Die unabhängige Variable war die Formulierungsvariante des Feedbacks mit zwei Ausprägungen, entweder *beibehalten vs. verändern* oder *gut vs. schlecht*. Die Formulierungsvariante wurde zwischen den Teilnehmenden variiert.

Als weitere unabhängige Variable wird im explorativen Teil die Selbstwertschätzung herangezogen, dafür wurden die Teilnehmenden anhand eines Mediansplits im Punktwert des Tests zur Selbstwertschätzung (Näheres hierzu in Kapitel 4.1.4) in zwei Bedingungen, d. h. eine mit hoher und eine mit niedriger Selbstwertschätzung, eingeteilt.

Als abhängige Variable wurde für die Hypothese H₁ die Einstellung zum Feedback als Differenz aus der Einstellung zum Feedback im aktuellen Fall (abgefragt nach der Abgabe des Feedbacks) und der generellen Einstellung zum Feedback (abgefragt vor der Abgabe des Feedbacks) erhoben. Bei Hypothese H₂ war die abhängige Variable die Einschätzung des Feedbacks durch den*die Teilnehmer*in als hilfreich. Die Hypothesen H₃ und H₄ beziehen sich auf die frei formulierten Verhaltensvorsätze der Teilnehmenden. In Hypothese H₃ wurde die Konkretheit der rückgemeldeten Aspekte untersucht. Die Konkretheit wurde nach dem Abstraktheitssuffix-Verfahren von Günther und Groeben (1978) kodiert. In Hypothese H₄ wurde die Anzahl der formulierten Verhaltensvorsätze nach einem durch die Autorin entwickelten Auszählschema als abhängige Variable herangezogen.

Details zur Kodierung werden in Kapitel 4.1.5 dargestellt.

Da vermutet wurde, dass die Leistung in der Bearbeitung der Aufgaben Auswirkungen auf die Akzeptanz des Feedbacks haben könnte, wurde diese miterhoben. Da bekannt ist, dass sich die selbst wahrgenommene von der objektiv erfassten Leistung unterscheidet (Shrauger, 1972) und in diesem Fall die subjektive Wahrnehmung der Leistung relevant ist, sollten die Teilnehmenden ihre Zufriedenheit mit der Leistung selbst angeben.

4.1.2 Untersuchungsablauf

Die Teilnehmenden nahmen in Gruppen von zwei bis vier Personen an der Studie teil. Bei jeder Durchführung waren zwei Beobachtende anwesend, die jeweils eine*n oder zwei Teilnehmende*n beobachteten und ihnen anschließend verhaltensbezogenes Feedback gaben. Die Beobachtenden waren Psychologiestudierende. Sie wurden in einer eintägigen Beobachtungs- und Feedbackschulung trainiert. Inhalte des Trainings waren die Vermittlung theoretischen Basiswissens zu beiden Themenfeldern sowie Übungen zu Wahrnehmungstendenzen. Die Verhaltenszuordnung zu den auf den Beobachtungsbögen (siehe Anhang E) und Beobachtungsinstruktionen (siehe Anhang F) vorgegebenen Dimensionen sowie ein Bezugsrahmentraining wurden anhand von Videobeispielen durchgeführt. Mit den Feedbackinstruktionen (siehe Anhang G) wurden die Beobachtenden in praktischen Übungen vertraut gemacht. Als Coverstory wurde ihnen mitgeteilt, dass untersucht werde, wie genau sie sich an verschiedene Beobachtungs- und Feedbackinstruktionen hielten und dass sie versuchen sollten, möglichst genau beim jeweiligen Schema zu bleiben.

Die jeweilige Gruppensitzung begann mit einer Begrüßung der Teilnehmenden, einem kurzen Überblick über den Ablauf sowie der Einholung von Datenschutz-Grundverordnungskonformen Einverständniserklärungen über Teilnahme, Verarbeitung der Daten und Veröffentlichung der anonymisierten Daten im Rahmen einer wissenschaftlichen Publikation (siehe Abbildung 5). Anschließend füllten die Teilnehmenden Fragebogen 1 zu soziodemographischen Angaben, der generellen Einstellung gegenüber dem Erhalten von Feedback und der Selbstwertschätzung aus (Näheres zu den Fragebögen in Kapitel 4.1.4).

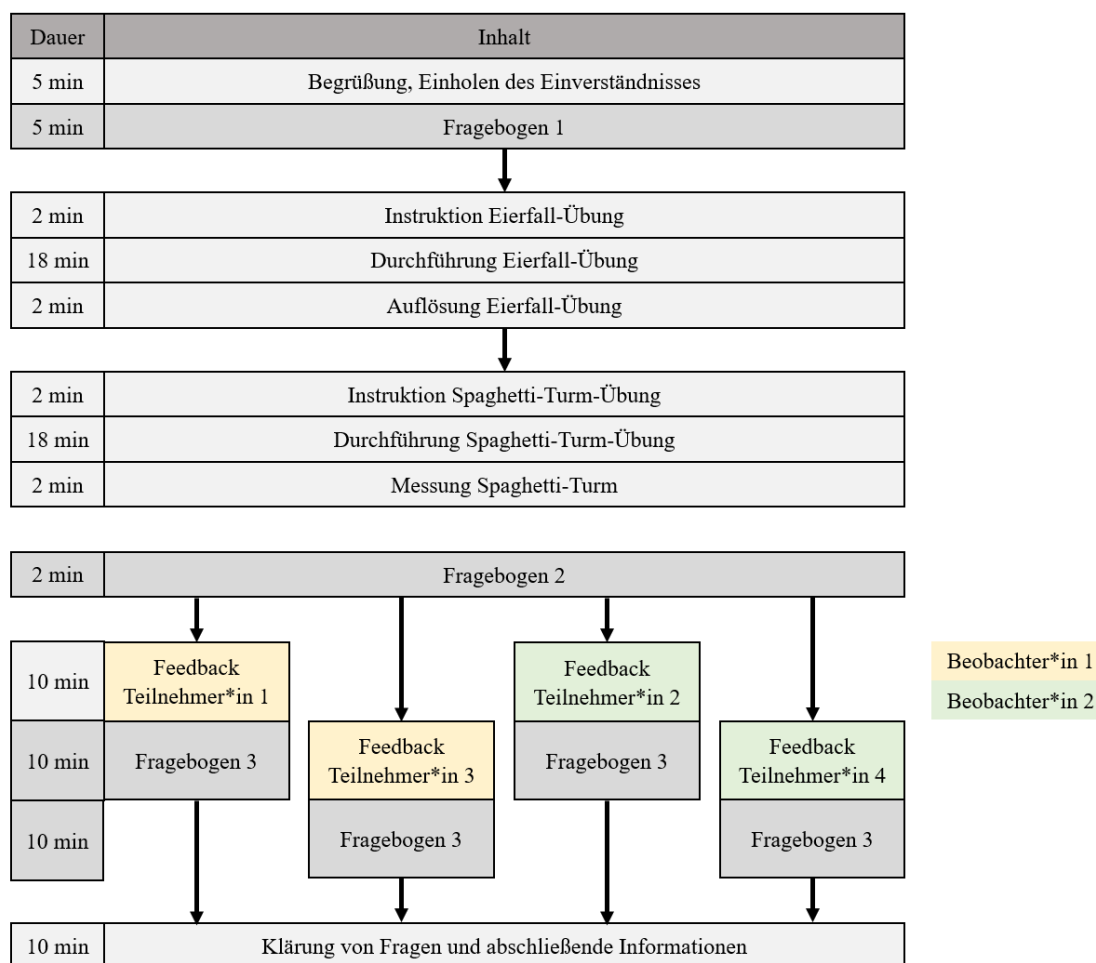


Abbildung 5. Zeitlicher Ablauf in Studie 2.

Es folgte die erste Kooperationsaufgabe, die Eierfall-Übung, in der die Teilnehmenden aus bestimmten bereitgestellten Materialien (siehe Anhang H) ein Ei so verpacken sollten, dass es einen Sturz aus zwei Metern Höhe unbeschadet überstehen kann. Beobachtet wurden Kooperation und Ziel- und Ergebnisorientierung, die Beobachtenden positionierten sich so im Raum, dass sie ihre jeweiligen Beobachtungsobjekte gut im Blick hatten, allerdings mit ausreichend Abstand. Die Teilnehmenden saßen gemeinsam an einem Tisch, auf dem Tisch die Materialien und Instruktionen (siehe Anhang I). Die Teilnehmenden wurden instruiert, dass während der Aufgabendurchführung keine Fragen möglich waren, die Beobachtenden würden keine Auskunft geben und die Studienleitung verließ den Raum und kam erst nach Fertigstellung der Aufgabe bzw. nach spätestens 20 Minuten zurück in den Raum. Nach Ablauf der Zeit wurde die Eierfall-Übung aufgelöst und das Ei aus zwei Metern Höhe sichtbar für die Teilnehmenden fallengelassen. Anschließend folgten Instruktionen und Materialbereitstellung für die Spaghetti-Turm-Übung (siehe Anhänge H und I). In dieser Übung

sollten die Teilnehmenden aus den vorhandenen Materialien – u. a. Spaghetti – einen möglichst hohen, freistehenden Turm bauen. Dieser sollte so stabil sein, dass er zum Zeitpunkt der Messung ein dem Material beigelegtes Marshmallow tragen konnte (der höchste Turm ist in Anhang J zu sehen). Es wurden die gleichen Dimensionen beobachtet wie in der Eierfall-Übung und die Studienleitung war wiederum während der Durchführung nicht im Raum. Grund für die Durchführung zweier Übungen war die Ausweitung der Beobachtungszeit. Nach 18 Minuten wurde die Höhe des Turms gemessen – insofern er freistehend war – und für die Gewinnfeststellung der als Teilnahmeanreiz versprochenen Amazon-Gutscheine (siehe Kapitel 4.1.3) notiert.

Nach der Messung schätzten die Teilnehmenden in Fragebogen 2 ihre Leistung in den zuvor bearbeiteten Aufgaben selbst ein. Darauf folgte das Feedback. Da jede*r Beobachtende bis zu zwei Teilnehmende beobachtet hatte, erhielt zunächst je ein*e Teilnehmer*in das Feedback in einem Vier-Augen-Gespräch, während der*die andere vor dem jeweiligen Raum wartete und das Feedback erst erhielt, nachdem das Gespräch mit der ersten Person beendet war. Direkt nach der Beendigung des Feedbacks füllten die Teilnehmenden Fragebogen 3 zur Einstellung gegenüber dem Erhalten von Feedback in diesem Fall ein; außerdem schätzten sie ein, wie hilfreich dieses Feedback war und trafen Verhaltensvorsätze.

Nach Abschluss aller Feedbackgespräche und Bearbeitung des letzten Fragebogens wurden offene Fragen zum Feedback sowie der Erhebung beantwortet. Die Teilnehmenden wurden darauf hingewiesen, dass sie jederzeit die Studienleitung kontaktieren könnten und erhielten bei Bedarf Kontaktdaten von Beratungsstellen. Die Teilnehmenden wurden gebeten, keine Informationen über die Erhebung an Dritte weiterzugeben.

Die Aufklärung über den wahren Zweck der Erhebung inkl. Nennung der Hypothesen erfolgte nach Abschluss der Erhebung per E-Mail.

4.1.3 Stichprobe

Insgesamt 70 Teilnehmende nahmen freiwillig an der Erhebung teil, davon 36 in der BV-Bedingung und 34 in der GS-Bedingung. Die meisten Teilnehmenden waren Studierende, wobei (internationale) Betriebswirtschaftslehre (mit 51.4%) und (Schul-)Psychologie (mit 25.7%) die am häufigsten vertretenen Studiengänge waren, zwei Teilnehmende waren in der Wissenschaft bzw. als Lehrkraft berufstätig. Das Alter der Teilnehmenden lag zwischen 19

und 33 Jahren. 60% der Teilnehmenden waren weiblich. Die Teilnehmenden wurden über die in Anhang K beigelegte Ausschreibung für eine Assessment-Center-Übung angeworben, die über soziale Netzwerke, den Presseverteiler der Otto-Friedrich-Universität Bamberg sowie E-Mail-Verteiler verschiedener Lehrveranstaltungen gestreut wurde. Die Teilnehmenden meldeten sich online für einen möglichen Termin an. An jedem Termin nahmen zwei bis vier Personen an einer Erhebung teil. 17 Teilnehmende erhielten für ihre Teilnahme Versuchspersonenstunden im Rahmen des Psychologie-Bachelorstudiums. Ein*e Teilnehmer*in erhielt eine Bestätigung über die Teilnahme als Beilage zu Bewerbungsunterlagen. Diejenigen Teilnehmenden, die die beste Leistung in einer der beiden Aufgaben zeigten, erhielten jeweils einen Amazon-Gutschein über 25 Euro (die Zweitbesten über 15 Euro, die Drittbesten über 10 Euro).

Bei einer teilnehmenden Person fehlte im Test zur Selbstwertschätzung ein Wert. Dieser wurde durch den Mittelwert der übrigen neun Werte derselben*desselben Teilnehmenden ergänzt.

Bei einer zweiten teilnehmenden Person fehlte die Angabe zur generellen Einstellung gegenüber Feedback. Wegen der ohnehin geringen Varianz dieser Variable und da nur ein Wert in allen Daten fehlte, wurde der fehlende Wert durch einen Mittelwert über alle Teilnehmenden in der Variable zur generellen Einstellung gegenüber Feedback ersetzt.

Für die Hypothese H_3 zur Konkretheit der Verhaltensvorsätze mussten drei Datensätze ausgeschlossen werden, da keine Verhaltensvorsätze getroffen wurden, deren Konkretheit hätte bewertet werden können. Für die übrigen Hypothesen wurden diese Datensätze nicht ausgeschlossen, da die Einstellung der Teilnehmenden zum Feedback sowie die Nützlichkeit des Feedbacks erhoben werden konnte und auch die Anzahl der Verhaltensvorsätze (nämlich keine) gewertet werden konnte.

Die Zuweisung zu den verschiedenen Experimentalbedingungen (BV-Bedingung oder GS-Bedingung) erfolgte randomisiert anhand von Zufallszahlen. An einem Durchführungstermin wurden jeweils alle Teilnehmenden entweder der BV-Bedingung oder der GS-Bedingung zugeordnet, um den Beobachtenden die Beobachtung und das Feedbackgeben zu erleichtern. Es ergab sich demnach ein Datensatz von $n_{\text{ges}} = 70$ [67]³, davon $n_{\text{BV}} = 36$ [35] und $n_{\text{GS}} = 34$ [32].

³ In eckigen Klammern stehen die analogen Kennwerte für die in Hypothese H_3 verwendete Stichprobe, insofern sich diese vom vollständigen Datensatz unterscheiden.

Die beiden Bedingungen unterschieden sich weder nennenswert hinsichtlich Alter der Teilnehmenden ($M_{BV} = 23.4$ [23.6] Jahre, $Mdn_{BV} = 23$ Jahre, $SD_{BV} = 3.4$ [3.3] Jahre; $M_{GS} = 23.9$ Jahre, $Mdn_{GS} = 24$ Jahre, $SD_{GS} = 3.3$ [3.4] Jahre), noch hinsichtlich des Studiengangs (BV-Bedingung: 52.8% [54.3%] (internationale) Betriebswirtschaftslehre, 27.8% [28.6%] (Schul-)Psychologie; GS-Bedingung: 50.0% [53.1%] (internationale) Betriebswirtschaftslehre, 23.5% [18.8%] (Schul-)Psychologie).

In der GS-Bedingung waren etwas mehr Frauen (BV-Bedingung: 52.8% [54.3%] weiblich, GS-Bedingung 70.6% [71.9%] weiblich), diese Verteilung war jedoch nicht signifikant $\chi^2(1) = 1.64, p = .199$ [$\chi^2(1) = 1.52, p = .217$].

Im Durchschnitt nahmen etwas mehr Teilnehmende in der BV-Bedingung ($M_{BV} = 3.22$, $SD_{BV} = 0.80$, $n_{BV} = 36$ [$M_{BV} = 3.23$, $SD_{BV} = 0.81$, $n_{BV} = 35$]) an einem Termin teil als in der GS-Bedingung ($M_{GS} = 3.00$, $SD_{GS} = 0.70$, $n_{GS} = 34$ [$M_{GS} = 3.00$, $SD_{GS} = 0.72$, $n_{GS} = 32$]). Dieser Unterschied war nicht signifikant, $t(67.61) = -1.24, p = .218, d = -0.30$ [$t(64.96) = -1.23, p = .225, d = -0.30$].

Die Teilnehmenden zeigten durchschnittlich in der BV-Bedingung ($M_{BV} = 5.58$, $SD_{BV} = 0.99$, $n_{BV} = 36$) eine etwas negativere generelle Einstellung zu Feedback als in der GS-Bedingung ($M_{GS} = 5.85$, $SD_{GS} = 1.02$, $n_{GS} = 34$). Dieser Unterschied war nicht signifikant, $t(67.54) = 1.15, p = .253, d = 0.28$. Auch wenn dementsprechend keine Differenzbildung zur Berechnung der abhängigen Variable in Hypothese H_1 notwendig wäre (weil kein statistisch signifikanter Unterschied in den Bedingungen besteht in der generellen Einstellung gegenüber Feedback), wird dennoch wie in Kapitel 4.1.1 beschrieben darauf zurückgegriffen, um bei der recht kleinen Stichprobe auch minimale Schwankungen zu berücksichtigen.

Darüber hinaus schätzten die Teilnehmenden in der BV-Bedingung ($M_{BV} = 4.33$, $SD_{BV} = 1.60$, $n_{BV} = 36$ [$M_{BV} = 4.31$, $SD_{BV} = 1.62$, $n_{BV} = 35$]) ihre Leistung schlechter ein als diejenigen in der GS-Bedingung ($M_{GS} = 5.47$, $SD_{GS} = 1.24$, $n_{GS} = 34$ [$M_{GS} = 5.53$, $SD_{GS} = 1.19$, $n_{GS} = 34$]). Dieser Unterschied war signifikant, $t(65.44) = 3.33, p = .001, d = 0.82$ [$t(62.19) = 3.52, p < .001, d = 0.89$], aber ist nicht experimentell bedingt – dies kann durch die in Kapitel 4.1.2 dargestellte Reihenfolge ausgeschlossen werden. In der Konsequenz wurde die Variable *selbsteingeschätzte Leistung* als Kovariate in die Analyse aufgenommen.

Im Durchschnitt war die Selbstwertschätzung bei Teilnehmenden in der BV-Bedingung ($M_{BV} = 23.56$, $SD_{BV} = 4.40$, $n_{BV} = 36$ [$M_{BV} = 23.54$, $SD_{BV} = 4.47$, $n_{BV} = 35$]) an einem Termin teil als in der GS-Bedingung ($M_{GS} = 23.08$, $SD_{GS} = 5.35$, $n_{GS} = 34$ [$M_{GS} = 23.24$, $SD_{GS} =$

5.22, $n_{GS} = 32$). Dieser Unterschied war nicht signifikant, $t(64.03) = -0.41$, $p = .684$, $d = -0.10$ [$t(61.33) = -0.26$, $p = .798$, $d = -0.07$].

Zur Verteilung der Beobachtenden auf die Bedingungen ist anzumerken, dass die Beobachtenden jeweils unterschiedlich häufig (im Vergleich der Beobachtenden miteinander, aber auch im Vergleich von Bedingung zu Bedingung) beobachteten und Feedback gaben. Die höchste bzw. niedrigste Anzahl an Beobachtungen inkl. Feedbackgeben je Beobachter*in betrug 10 bzw. 4 – die anderen Beobachtenden lagen dazwischen ($M = 7.00$, $SD = 2.05$ [$M = 6.70$, $SD = 1.95$]). Zwei [drei] Beobachtende waren genauso häufig in der BV-Bedingung wie in der GS-Bedingung im Einsatz, die maximale Diskrepanz betrug in absoluten Zahlen 4, d. h., ein*e Beobachter*in war vier Mal häufiger in der BV-Bedingung als in der GS-Bedingung im Einsatz.

4.1.4 Fragebögen

Die Fragebögen waren in der BV- und der GS-Bedingung identisch.

Fragebogen 1 (siehe Anhang L) erfasste zunächst soziodemographische Variablen wie Geschlecht, Studiengang und Alter. Anschließend wurde auf einer siebenstufigen Ratingskala (1 = *überhaupt nicht*, 7 = *voll und ganz*) die Zustimmung zur Aussage „Ich bekomme generell gerne Feedback“ erhoben. Dann folgte die Rosenberg Selbstwertschätzungsskala in der revidierten Fassung von Collani und Herzberg (2003). Die zehn Items sind jeweils auf einer vierstufigen Ratingskala (1 = *trifft gar nicht zu* bis 4 = *trifft voll und ganz zu*) zu bewerten.

In Fragebogen 2 (siehe Anhang M) gaben die Teilnehmenden auf einer siebenstufigen Ratingskala (1 = *überhaupt* bis 7 = *voll und ganz*) an, wie sehr sie der Aussage „Ich bin mit meiner Leistung zufrieden“ zustimmen.

In Fragebogen 3 (siehe Anhang N) gaben die Teilnehmenden zunächst an, wie sehr sie den Aussagen „Ich habe dieses Mal gerne Feedback bekommen“ und „Dieses Feedback war hilfreich“ auf einer siebenstufigen Ratingskala (1 = *überhaupt* bis 7 = *voll und ganz*) zustimmen. In einem Freitextfeld darunter erhielten die Teilnehmenden bei beiden Aussagen die Möglichkeit, Gründe für ihre Angabe zu erläutern. Dies diente als zusätzliche Informationsquelle für den Fall, dass Ausreißer aufgetreten wären. Anschließend beantworteten die Teilnehmenden im Freitextfeld unter der Überschrift *Meine Vorsätze* folgende Instruktion „Damit

du aus der heutigen Übung etwas mitnehmen kannst, solltest du dir im letzten Teil des Fragebogens einige Vorsätze notieren. Was möchtest du beim nächsten Mal anders machen? Welche Erkenntnisse nimmst du mit?“. Die Fragen wurden bewusst offen gewählt, sodass die Formulierungen und der Umfang für die Teilnehmenden frei wählbar waren (dies wurde als abhängige Variable erhoben).

4.1.5 Kodierung

Für die Hypothese H_3 zur Konkretheit der Vorsätze wurden die Rückmeldungen nach dem Abstraktheitssuffix-Verfahren von Günther und Groeben (1978) für deutsche Begriffe sowie nach dem analogen englischsprachigen Verfahren von Gillie (1957) für englische Begriffe ausgewertet. Diese Kodierung erfolgte analog zur Kodierung der Hypothese H_3 zur Konkretheit der Rückmeldungen in Studie 1, für eine detaillierte Darstellung sei daher auf Kapitel 3.1.4 verwiesen. Wie ebenda erwähnt, empfehlen Günther und Groeben (1978) das Verfahren ab 400 Substantiven. In Studie 2 lag die Anzahl der Substantive pro Teilnehmenden zwischen 1 und 15 ($M = 6.15$, $SD = 3.56$). Das Verfahren wurde dennoch herangezogen, da in dieser ersten Annäherung an die Thematik auf ein bewährtes Verfahren zurückgegriffen werden sollte. Die Interraterreliabilität, bestimmt über Cohens κ , für dieses Kodierverfahren lag in der hier vorliegenden Arbeit nach Doppelkodierung von 10% des Materials bei $\kappa = .88$, was nach Landis und Koch (1977) einer nahezu vollkommenen bzw. nach Greve et al. (1997) einer ausgezeichneten Übereinstimmung entspricht.

Analog zur Kodierung der Hypothese H_3 in Studie 1 wurde die über dieses Verfahren bestimmte Anzahl an Abstraktheitssubstantiven wurde von der Gesamtzahl der verwendeten Substantive pro Teilnehmer*in subtrahiert und dieser Wert an der Gesamtzahl der verwendeten Substantive pro Fragebogen relativiert. So ergab sich die Konkretheit der Vorsätze eines*r Teilnehmenden.

Für die Hypothese H_4 zur Anzahl der Vorsätze musste zunächst festgestellt werden, welche Textpassagen im Freitextfeld als Vorsatz zu werten sind. Dies war nötig, da durch die offene Formulierung der Instruktion inhaltlich verschiedene Beiträge zu verzeichnen waren. Als Vorsatz wurde demnach alles gewertet, was beschreibt, dass die Person selbst eine Verhaltensweise oder Eigenschaft zukünftig beibehalten oder verändern will – dabei müssen nicht alle Kriterien erfüllt sein, aber es darf diesen zumindest nicht offensichtlich widersprechen. Anschließend wurden die Vorsätze ausgezählt. Dabei war zu berücksichtigen, dass auch

mehrere Vorsätze in einem Satz enthalten sein konnten, insofern sie sich inhaltlich unterscheiden. Ob ein inhaltlicher Unterschied bestand, wurde im Zweifelsfall anhand der Synonymanzeige auf duden.de bestimmt. Die Anzahl der nicht synonymen Vorsätze wurde dann je Teilnehmer*in bestimmt.

Die Kodierregeln sind im Detail in Anhang O zu finden.

4.2 Ergebnisse

4.2.1 Einstellung zum Feedback

In Hypothese H₁ wurde angenommen, dass sich in der BV-Bedingung eine positivere Einstellung zum Feedback zeigt. Wie in den Kapiteln 4.1.1 und 4.1.3 erläutert, wurde hierzu die Differenz aus der Einstellung zum aktuellen Feedback und der generellen Einstellung zum Feedback herangezogen. In ebendiesen Kapiteln wurde darüber hinaus dargestellt, dass die Leistung als Kovariate in die statistische Überprüfung mit aufgenommen wird.

Voraussetzungsprüfung für Kovarianzanalyse (nach Bortz & Schuster, 2016). Für die Durchführung einer ANCOVA sind zunächst die Voraussetzungen einer ANOVA zu prüfen. Die Unabhängigkeit der Messungen ist realisiert durch die verschiedenen Bedingungen, zwischen denen keine Überlappung besteht. Die Normalverteilungsannahme für die Einstellungsdifferenz gegenüber dem Feedback ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.87$ und $p_{BV} < .001$ sowie $W_{GS} = 0.88$ und $p_{GS} = .002$. Da der F-Test jedoch gegenüber dieser Annahmeverletzung bei einer annähernd gleich großen Bedingungsbesetzung wie in diesem Fall ($n_{BV} = 36$, $n_{GS} = 34$) robust ist, kann er trotzdem herangezogen werden (Glass, Peckham & Sanders, 1972; Harwell, Rubinstein, Hayes & Olds, 2016; Lix, Keselman & Keselman, 1996). Die Varianzhomogenität nach dem Levene-Test ist gegeben, $F(1, 68) = 1.44$, $p = .235$.

Darüber hinaus werden zwei Voraussetzungen spezifisch für die ANCOVA als relevant angesehen. Häufig wird die Unabhängigkeit der Kovariate vom Effekt der experimentellen Bedingung als wichtige Voraussetzung angeführt (Miller & Chapman, 2001), dies ist jedoch umstritten (Overall & Woodward, 1977a; Overall & Woodward, 1977b). Gleichwohl wurde dies in dem vorliegenden Experimentaldesign ausreichend berücksichtigt (T. Lösch, persönliche Kommunikation, 25. Juni 2018), da die Kovariate *Leistung* nicht infolge der experimentellen Manipulation schwankte (diese erfolgte erst danach), sondern sich zufällig

zwischen den Bedingungen unterschied (siehe Kapitel 4.1.3). Als zweite ANCOVA-spezifische Voraussetzung sind die homogenen Regressionen innerhalb der Stichproben zu nennen, die zwar bei annähernd gleich großen Anzahlen in den Bedingungen wie in diesem Fall gegeben ($n_{BV} = 36$, $n_{GS} = 34$) nicht zum Tragen kommen (Hamilton, 1977; Overall & Woodward, 1977a), aber der Vollständigkeit halber überprüft wurden, indem die Interaktion zwischen *Feedbackvariante* und der Kovariate *Leistung* in Bezug auf die abhängige Variable *Einstellungsdifferenz* auf Signifikanz überprüft wurde (Field, Miles & Field, 2012). Diese Interaktion war nicht signifikant, $F(1, 66) < 0.01$, $p = .994$, $\eta_p^2 < .01$, die Homogenität der Innerhalb-Regressionen ist demnach gegeben.

Die Durchführung einer ANCOVA ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel gaben die Teilnehmenden im aktuellen Fall – bereinigt um ihre generelle Einstellung zum Feedback – in der BV-Bedingung ($M_{BV} = 1.20$, $SD_{BV} = 0.95$, $n_{BV} = 36$) lieber Feedback als in der GS-Bedingung ($M_{GS} = 0.65$, $SD_{GS} = 1.32$, $n_{GS} = 34$), siehe Tabelle 2. Der Haupteffekt für die Feedbackvariante ist signifikant, $F(1, 67) = 4.31$, $p = .042$, $\eta_p^2 = .06$. Die Kovariate *Leistung* wird nicht signifikant, $F(1, 67) = 0.30$, $p = .584$, $\eta_p^2 < .01$.

Die Hypothese H_1 , dass sich in der BV-Bedingung eine positivere Einstellung zum Feedback zeigt, kann demnach beibehalten werden.

Tabelle 2. Mittelwert und Standardabweichung der Einstellung zum Feedbacknehmen generell, im aktuellen Fall und als Differenz

Variable	Untersuchungsbedingung			
	BV-Bedingung		GS-Bedingung	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Generell ^a	5.58	0.99	5.85	1.02
Aktuell ^b	6.78	0.42	6.50	1.05
Differenz ^c	1.20	0.95	0.65	1.32

Anmerkungen. Die Zustimmung zum jeweiligen Item wurde auf einer Ratingskala erhoben (1 = überhaupt nicht, 7 = voll und ganz), $n_{BV} = 36$ und $n_{GS} = 34$.

^a Das Item lautete „Ich bekomme generell gerne Feedback“. ^b Das Item lautete „Ich habe dieses Mal gerne Feedback bekommen“. ^c Berechnet wurde je Teilnehmer*in die Einstellung zum aktuellen Feedback abzüglich der Einstellung zum Feedback generell. ^d BV-Bedingung steht für die Bedingung, in der die Teilnehmenden Feedback in der Variante *beibehalten vs. verändern* erhalten. ^e GS-Bedingung steht für die Bedingung, in der die Teilnehmenden Feedback in der Variante *gut vs. schlecht* erhalten.

Explorative Prüfungen. Bei Betrachtung der Mittelwerte in Tabelle 2 ist auffällig, dass im aktuellen Fall im Schnitt deutlich lieber Feedback gegeben wurde als generell. Dieser Unterschied ist sowohl in der BV-Bedingung auf einem nach Bonferroni-Korrektur adjustiertem Alphafehlerniveau von $\alpha = .05/2 = .025$ signifikant, $t(35) = 7.59$, $p < .001$, $d = 2.57$, als auch in der GS-Bedingung, $t(33) = 2.85$, $p = .007$, $d = 0.99$. Die Alphafehler-Korrektur wurde in dieser Testung vorgenommen, da die gleiche abhängige Variable in zwei Tests (einmal innerhalb der BV- und einmal innerhalb der GS-Bedingung) inferenzstatistisch überprüft wurde.

4.2.2 Nützlichkeit des Feedbacks

In Hypothese H₂ wurde angenommen, dass in der BV-Bedingung das Feedback als hilfreicher empfunden wird. Wie in den Kapiteln 4.1.1 und 4.1.4 erläutert, wurde hierzu die Einschätzung des Feedbacks auf einer siebenstufigen Ratingskala herangezogen. Im erstgenannten Kapitel wurde darüber hinaus dargestellt, dass die Leistung als Kovariate in die statistische Überprüfung mit aufgenommen wird.

Voraussetzungsprüfung für Kovarianzanalyse (nach Bortz & Schuster, 2016). Analog zu Hypothese H₁ sind auch für die H₂ die Voraussetzungen zu prüfen. Der Einfachheit halber

wird an dieser Stelle auf ausführliche Erläuterungen verzichtet, stattdessen werden die statistischen Kennwerte sowie die Erfüllung der Voraussetzungen knapp berichtet (für eine ausführliche Darstellung vgl. Kapitel 4.2.1). Zunächst werden die Voraussetzungen der ANOVA geprüft: Die Unabhängigkeit der Messungen ist gegeben. Die Normalverteilungsannahme für die Nützlichkeit des Feedbacks ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.75$ und $p_{BV} < .001$ sowie $W_{GS} = 0.71$ und $p_{GS} < .001$, was jedoch einer Nutzung des Tests nicht widerspricht. Die Varianzhomogenität nach dem Levene-Test ist gegeben, $F(1, 68) = 0.51$, $p = .479$. Darüber hinaus werden die Voraussetzungen für die ANCOVA geprüft: Die Unabhängigkeit der Kovariate vom Effekt der experimentellen Bedingungen wird als gegeben angesehen, die Homogenität der Innerhalb-Regressionen ist nicht verletzt, $F(1, 66) < 0.01$, $p = .994$, $\eta^2 < .01$.

Die Durchführung einer ANCOVA ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel empfanden die Teilnehmenden das Feedback in der BV-Bedingung ($M_{BV} = 5.97$, $SD_{BV} = 1.36$, $n_{BV} = 36$) als etwas weniger hilfreich als in der GS-Bedingung ($M_{GS} = 6.21$, $SD_{GS} = 1.15$, $n_{GS} = 34$). Signifikant ist weder der Haupteffekt für die *Feedbackvariante*, $F(1, 67) = 0.01$, $p = .943$, $\eta^2 < .01$, noch der Effekt der Kovariate *Leistung*, $F(1, 67) = 3.12$, $p = .082$, $\eta^2 = .04$. Eine Sensitivitätsanalyse ergab, dass unter den gegebenen Parametern und Stichprobengrößen der Effekt $\eta^2 \geq .11$ hätte sein müssen, um ihn – insofern er in der Population tatsächlich existiert – mit $> 80.0\%$ auch vorfinden zu können. Die Teststärkebestimmung a posteriori ergab bei der errechneten Effektgröße der Feedbackvariante eine Teststärke von 5.1 %.

Die Hypothese H_2 , dass das Feedback in der BV-Bedingung als hilfreicher empfunden wird, muss verworfen werden.

Explorative Prüfungen. Auffällig sind die insgesamt recht hohen Mittelwerte, die auf einen Deckeneffekt hindeuten. In Abbildung 6 sind nach Bedingungen die prozentualen Häufigkeiten der einzelnen Ratingangaben dargestellt. In beiden Bedingungen haben mindestens $\frac{3}{4}$ der Teilnehmenden entweder den Wert 6 oder 7 auf der Ratingskala als Einschätzung dessen angegeben, für wie hilfreich sie dieses Feedback hielten.

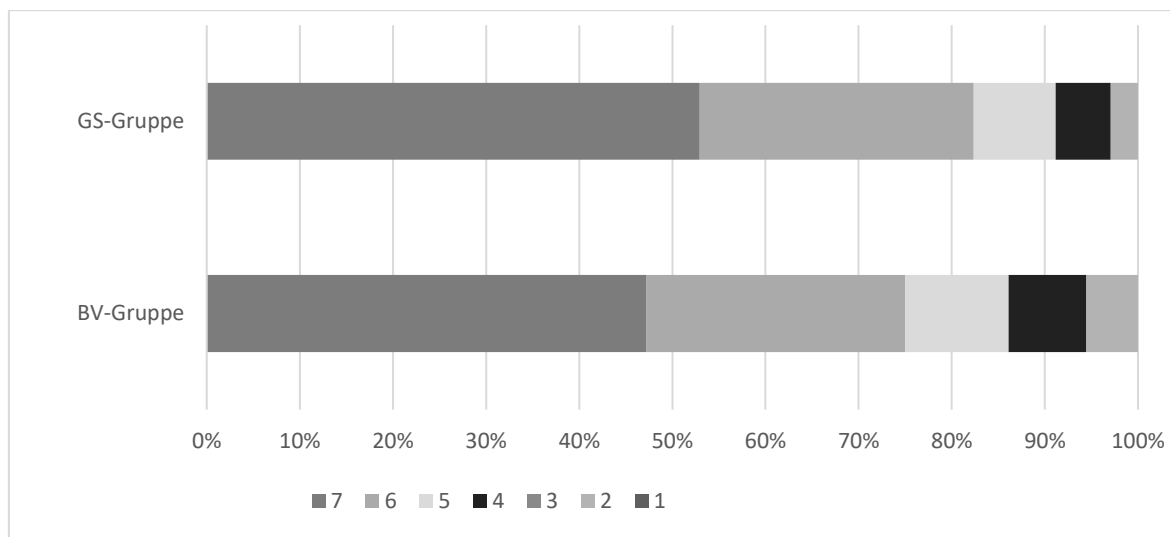


Abbildung 6. Prozentuale Häufigkeit der Ratingangaben von 1 = *überhaupt nicht* bis 7 = *voll und ganz*.

4.2.3 Konkretheit der Vorsätze

In Hypothese H₃ wurde angenommen, dass in der BV-Bedingung die Vorsätze konkreter sind. Wie in den Kapiteln 4.1.1 und 4.1.5 erläutert, wurde die Anzahl an Abstraktheitssubstantiven (nach Günther & Groeben, 1978) von der Anzahl aller Substantive subtrahiert und dieser Wert der konkreten Substantive an der Anzahl aller Substantive relativiert. Dies ergab die Konkretheitsrate je Teilnehmer*in. Im erstgenannten Kapitel wurde darüber hinaus dargestellt, dass die Leistung als Kovariate in die statistische Überprüfung mit aufgenommen wird.

Voraussetzungsprüfung für Kovarianzanalyse (nach Bortz & Schuster, 2016). Analog zu den Hypothese H₁ und H₂ sind auch für die H₃ die Voraussetzungen zu prüfen. Der Einfachheit halber wird an dieser Stelle auf ausführliche Erläuterungen verzichtet, stattdessen werden die statistischen Kennwerte sowie die Erfüllung der Voraussetzungen knapp berichtet (für eine ausführliche Darstellung vgl. Kapitel 4.2.1). Zunächst werden die Voraussetzungen der ANOVA geprüft: Die Unabhängigkeit der Messungen ist gegeben. Die Normalverteilungsannahme für die Konkretheit der Vorsätze ist in beiden Bedingungen nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.93$ und $p_{BV} = .025$ sowie $W_{GS} = 0.88$ und $p_{GS} = .002$, was jedoch einer Nutzung des Tests nicht widerspricht. Die Varianzhomogenität nach dem Levene-Test ist nicht gegeben, $F(1, 65) = 6.33$, $p = .014$, dank der annähernd gleich großen Bedingungsbesetzung stellt aber auch dies kein Problem dar (Bortz &

Schuster, 2016). Darüber hinaus werden die Voraussetzungen für die ANCOVA geprüft: Die Unabhängigkeit der Kovariate vom Effekt der experimentellen Bedingungen wird als gegeben angesehen, die Homogenität der Innerhalb-Regressionen ist nicht verletzt, $F(1, 63) = 0.08, p = .773, \eta_p^2 < .01$.

Die Durchführung einer ANCOVA ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel war die Konkretheitsrate in der BV-Bedingung ($M_{BV} = 0.79, SD_{BV} = 0.17, n_{BV} = 35$) etwas höher als in der GS-Bedingung ($M_{GS} = 0.73, SD_{GS} = 0.27, n_{GS} = 32$). Signifikant ist weder der Haupteffekt für die Feedbackvariante, $F(1, 64) = 0.41, p = .525, \eta_p^2 < .01$, noch der Effekt der Kovariate *Leistung*, $F(1, 64) = 0.92, p = .340, \eta_p^2 = .01$. Eine Sensitivitätsanalyse ergab, dass unter den gegebenen Parametern und Stichprobengrößen der Effekt $\eta_p^2 \geq .11$ hätte sein müssen, um ihn – insofern er in der Population tatsächlich existiert – mit $> 80.0\%$ auch vorfinden zu können. Die Teststärkebestimmung a posteriori ergab bei der errechneten Effektgröße der Feedbackvariante eine Teststärke von 9.9 %.

Die Hypothese H_3 , dass in der BV-Bedingung die Vorsätze konkreter sind, muss demnach verworfen werden.

4.2.4 Anzahl der Vorsätze

In Hypothese H_4 wurde angenommen, dass in der BV-Bedingung mehr Vorsätze getroffen werden als in der GS-Bedingung. Wie in Kapitel 4.1.5 erläutert, wurde hierzu die Anzahl der Vorsätze anhand zuvor definierter Regeln ausgezählt. In den Kapiteln 4.1.1 und 4.1.3 wurde darüber hinaus dargestellt, dass die Leistung als Kovariate in die statistische Überprüfung mit aufgenommen wird.

Voraussetzungsprüfung für Kovarianzanalyse (nach Bortz & Schuster, 2016). Analog zu den Hypothesen H_1, H_2 und H_3 sind auch für die H_4 die Voraussetzungen zu prüfen. Der Einfachheit halber wird an dieser Stelle auf ausführliche Erläuterungen verzichtet, stattdessen werden die statistischen Kennwerte sowie die Erfüllung der Voraussetzungen knapp berichtet (für eine ausführliche Darstellung vgl. Kapitel 4.2.1). Zunächst werden die Voraussetzungen der ANOVA geprüft: Die Unabhängigkeit der Messungen ist gegeben. Die Normalverteilungsannahme für die Anzahl der Vorsätze ist in der BV-Bedingung nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.84$ und $p_{BV} < .001$, in der GS-Bedingung

nicht, $W_{GS} = 0.94$ und $p_{GS} = .07$, was jedoch einer Nutzung des Tests nicht widerspricht. Die Varianzhomogenität nach dem Levene-Test ist gegeben, $F(1, 68) = 0.22, p = .638$. Darüber hinaus werden die Voraussetzungen für die ANCOVA geprüft: Die Unabhängigkeit der Kovariate vom Effekt der experimentellen Bedingungen wird als gegeben angesehen, die Homogenität der Innerhalb-Regressionen ist nicht verletzt, $F(1, 66) = 0.14, p = .705, \eta_p^2 < .01$.

Die Durchführung einer ANCOVA ist demnach zulässig.

Ergebnisse der Hypothesenprüfung. Im Mittel wurden etwas mehr Vorsätze in der BV-Bedingung ($M_{BV} = 3.72, SD_{BV} = 2.61, n_{BV} = 36$) gefasst als in der GS-Bedingung ($M_{GS} = 3.00, SD_{GS} = 1.95, n_{GS} = 34$). Signifikant ist weder der Haupteffekt für die Feedbackvariante, $F(1, 67) = 1.62, p = .208, \eta_p^2 = .02$, noch der Effekt der Kovariate *Leistung*, $F(1, 67) = 0.04, p = .848, \eta_p^2 < .01$. Eine Sensitivitätsanalyse ergab, dass unter den gegebenen Parametern und Stichprobengrößen der Effekt $\eta_p^2 \geq .11$ hätte sein müssen, um ihn – insofern er in der Population tatsächlich existiert – mit $> 80.0\%$ auch vorfinden zu können. Die Teststärkebestimmung a posteriori ergab bei der errechneten Effektgröße der Feedbackvariante eine Teststärke von 10.1 %.

Die Hypothese H_4 , dass in der BV-Bedingung mehr Vorsätze gefasst werden, muss demnach verworfen werden.

4.2.5 Explorative Untersuchungen zur Selbstwertschätzung

Im Folgenden werden sämtliche Hypothesentests aus Studie 2 unter Hinzunahme der Variable Selbstwertschätzung erneut gerechnet.

Datenaufbereitung. Anhand des Medians ($Mdn_{SWS} = 24$) in der Skala zur Selbstwertschätzung von Collani und Herzberg (2003) wurden die Teilnehmenden in die Gruppen hohe (für alle Punktwerte ≤ 24 ; $n_{hSWS} = 33$) vs. niedrige (für alle Punktwerte > 24 ; $n_{nSWS} = 37$) Selbstwertschätzung aufgeteilt. Wohlwissend, dass mit der Dichotomisierung ein Verlust an Teststärke einher geht (Cohen, 1983), wurde dieses Verfahren in dieser ersten explorativen Untersuchung zur Thematik vereinfachend herangezogen.

Kovarianzanalysen für die Hypothesen H₁, H₂ und H₃. Für die Hypothesen H₁ zur Einstellung gegenüber dem Feedback, H₂ zur Nützlichkeit des Feedbacks und H₃ zur Konkretheit der Vorsätze zeigten sich weder signifikante Haupteffekte noch Interaktionen. Diese werden deshalb nicht näher dargestellt.

Kovarianzanalyse für die Hypothese H₄. Analog zu den Prüfungen der Hypothesen ohne Hinzunahme der Selbstwertschätzung sind auch unter Hinzunahme der Selbstwertschätzung die Voraussetzungen für die Kovarianzanalyse im Rahmen von Hypothese H₄ zu prüfen. Der Einfachheit halber wird an dieser Stelle auf ausführliche Erläuterungen verzichtet, stattdessen werden die statistischen Kennwerte sowie die Erfüllung der Voraussetzungen knapp berichtet (für eine ausführliche Darstellung vgl. Kapitel 4.2.1). Zunächst werden die Voraussetzungen der ANOVA geprüft: Die Unabhängigkeit der Messungen ist gegeben. Die Normalverteilungsannahme für die Anzahl der Vorsätze ist in der BV-Bedingung nach dem Shapiro-Wilk-Test signifikant verletzt, $W_{BV} = 0.84$ und $p_{BV} < .001$, in der GS-Bedingung nicht, $W_{GS} = 0.94$ und $p_{GS} = .07$. In der Gruppe der niedrigen Selbstwertschätzung ist die Normalverteilungsannahme nach dem Shapiro-Wilk-Test nicht verletzt, $W_{nSWS} = 0.97$ und $p_{nSWS} = .294$, in der Gruppe mit hoher Selbstwertschätzung hingegen ist sie verletzt, $W_{hSWS} = 0.78$ und $p_{hSWS} < .001$. Dies widerspricht jedoch einer Nutzung des Tests nicht (Glass et al., 1972; Harwell et al., 2016; Lix et al., 1996), wenn die Bedingungs- bzw. Gruppenbesetzung annähernd gleich groß ist ($n_{BV} = 36$, $n_{GS} = 34$ bzw. $n_{nSWS} = 37$, $n_{hSWS} = 33$). Die Varianzhomogenität nach dem Levene-Test ist sowohl in der Aufteilung nach der *Selbstwertschätzung* gegeben, $F(1, 68) = 0.27$, $p = .608$, als auch in der Aufteilung nach der *Feedbackvariante*, $F(1, 68) = 0.22$, $p = .638$. Darüber hinaus werden die Voraussetzungen für die ANCOVA geprüft: Die Unabhängigkeit der Kovariate vom Effekt der experimentellen Bedingungen wird als gegeben angesehen, die Homogenität der Innerhalb-Regressionen ist nicht verletzt $F(1, 62) < 1.77$, $p > .21$, $\eta_p^2 < .03$.

Die Durchführung einer ANCOVA ist demnach wohl noch zulässig – wenn auch zu diskutieren ist, ob bei Stichprobengrößen wie hier vorliegend ($n_{nSWS} = 37$, $n_{hSWS} = 33$) noch von gleich großen Stichproben ausgegangen werden kann.

Ergebnisse der Hypothesenprüfung. Für Hypothese H₄ zur Anzahl der gefassten Vorsätze zeigt sich ein signifikanter Haupteffekt für die Selbstwertschätzungs-Gruppe, $F(1, 65) = 5.59$, $p = .021$, $\eta_p^2 = .08$, und eine signifikante Interaktion zwischen Selbstwertschätzungs-Gruppe und der Feedbackvariante, $F(1, 65) = 8.99$, $p = .004$, $\eta_p^2 = .12$.

Im Mittel fassten die Teilnehmenden mit niedriger Selbstwertschätzung ($M_{\text{nSWS}} = 3.46$, $SD_{\text{nSWS}} = 1.89$, $n_{\text{nSWS}} = 37$) mehr Vorsätze als die Teilnehmenden mit hoher Selbstwertschätzung ($M_{\text{hSWS}} = 3.27$, $SD_{\text{hSWS}} = 2.76$, $n_{\text{hSWS}} = 33$), siehe Abbildung 7. Unter den Teilnehmenden mit hoher Selbstwertschätzung wurden wiederum in der BV-Bedingung mehr Vorsätze als in der GS-Bedingung gefasst, für die Teilnehmenden mit niedriger Selbstwertschätzung zeichnete sich das gegenteilige Bild ab.

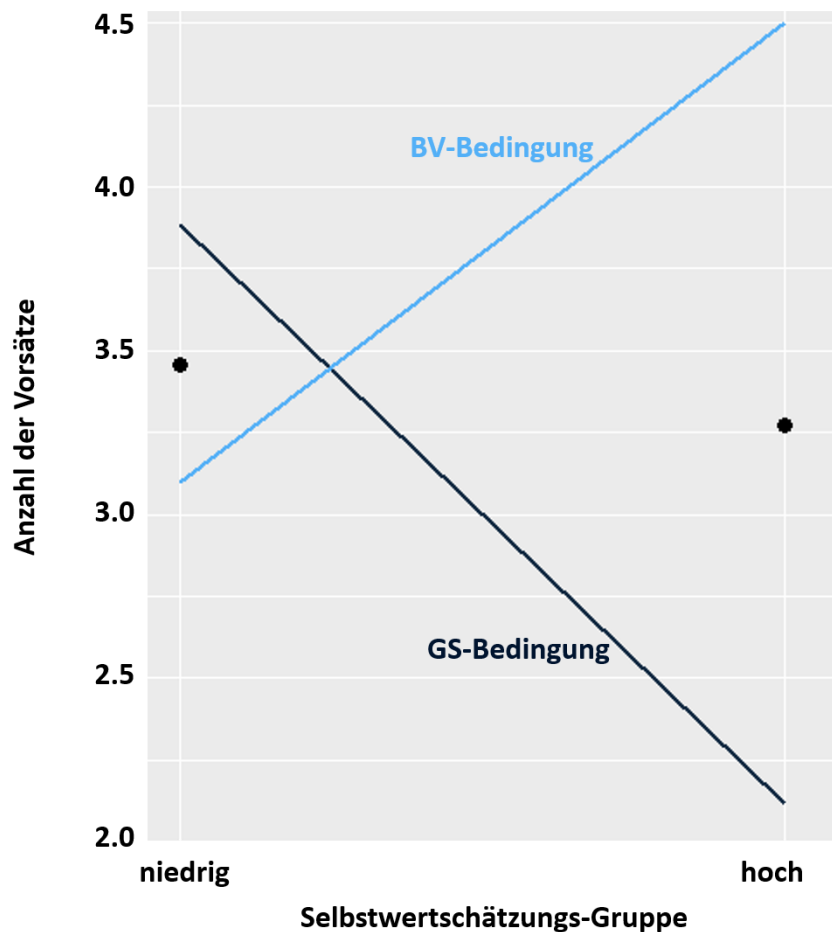


Abbildung 7. Anzahl der Vorsätze nach Selbstwertschätzungs-Gruppe und Feedbackvariante.

4.3 Diskussion

4.3.1 Diskussion der Hypothesen

In Studie 2 wurde untersucht, ob die Formulierung von Feedback *beibehalten vs. verändern* für Feedbacknehmende annehmbarer, hilfreicher und verhaltensänderungsrelevanter ist als die Formulierung *gut vs. schlecht*. Die erste Hypothese ließ sich bestätigen– die Einstellung zum Feedback war in der BV-Bedingung signifikant positiver als in der GS-Bedingung.

Im Folgenden sollen zunächst die Ergebnisse der Hypothesenprüfung für diese Studie einzeln diskutiert, in den aktuellen Forschungsstand eingeordnet und Ansatzpunkte für zukünftige Forschung dargestellt werden. In Kapitel 5 werden anschließend die Ergebnisse beider Studien zusammengeführt und ein Gesamtausblick gegeben.

Die in Hypothese H₁ angenommene positivere Einstellung zum Feedback in der BV-Bedingung ließ sich bestätigen, es zeigte sich ein signifikanter Unterschied. Eine auf diese Art und Weise operationalisierte geringere Belastung der*des Feedbacknehmenden (Besieux, 2017; Donnelly & Kirk, 2015) durch das Formulierungskonzept *beibehalten vs. verändern* im Gegensatz zu *gut vs. schlecht* ist weiterhin anzunehmen. Darüber hinaus zeigte sich, dass in dieser Studie die Teilnehmenden eine signifikant positivere Einstellung zum aktuellen Feedback hatten als generell – unabhängig zur Feedbackvariante. Dies geht einher mit Befunden zur geringen Annehmbarkeit und negativen Konsequenzen von Personfeedback (Brummelman, Thomaes, Overbeek et al., 2014; Cimpian et al., 2007; Corpus et al., 2006; Kamins & Dweck, 1999; Mueller & Dweck, 1998) und ist möglicherweise darauf zurückzuführen, dass prozessbezogenes Feedback von geschulten Beobachtenden auf Basis eines verhaltensverankerten Beobachtungsbogens gegeben wurde.

Die Annahme in Hypothese H₂, dass Feedback in der BV-Bedingung als hilfreicher empfunden wird, ließ sich nicht bestätigen. Die geringe Teststärke – die im Übrigen für alle geprüften Hypothesen in Studie 2 zu finden ist – ist in diesem Fall besonders kritisch zu betrachten. Unter gegebenen Parametern und Stichprobengrößen lag die Wahrscheinlichkeit eines signifikanten Testergebnisses bei tatsächlicher Gültigkeit der postulierten Hypothese bei nur 5.1 %. Für zukünftige Studien sollte eine den zu erwartenden Effekten angemessene Stichprobengröße untersucht werden. Darüber hinaus sind die recht hohen Mittelwerte, die in Richtung eines Deckeneffekts deuten, kritisch zu hinterfragen. Die Reduktion der Varianz, die mit einem Deckeneffekt einhergeht, führt zu einer verzerrten Schätzung der Daten in arithmetischem Mittel und Standardabweichung (Döring & Bortz, 2016).

Die Prüfung der Hypothese H₃ zeigte, dass in der Konkretheit der Vorsätze kein signifikanter Unterschied zwischen den beiden Bedingungen bestand. Die Annahme, dass das Formulierungskonzept nach *beibehalten vs. verändern* mehr Verhaltensvorsätze (als Indikator für tatsächliche Verhaltensänderung) generiert, kann nicht aufrechterhalten werden. Unter gegebenen Parametern und Stichprobengrößen lag die Teststärke bei dieser Hypothese bei 9.9%. Kritisch zu bewerten ist überdies die Verwendung des Abstraktheitssuffix-Verfahrens

(Günther & Groeben, 1978) zur Kodierung der Konkretheit der rückgemeldeten Aspekte. Dieses Verfahren ist ausgelegt für eine Mindestanzahl von 400 Substantiven pro Text, wobei die in dieser Studie zu bewertenden Textstücke im Mittel etwa sechs Substantive enthielten (siehe Kapitel 4.1.5). Dieses Verfahren wurde nach intensiver Prüfung von Alternativen dennoch herangezogen, da es ein standardisiertes, reliables und valides Maß (bei Berücksichtigung der Anwendungsbedingungen, die hier nur bedingt gegeben waren) darstellt. In zukünftiger Forschung zur Thematik ist dies zu berücksichtigen und ein zutreffenderes Maß auszuwählen bzw. zu entwickeln.

In Hypothese H₄ wurde geprüft, ob im Rahmen des neuen Formulierungskonzeptes die Anzahl gefasster Verhaltensvorsätze größer ist. Der hypothesenkonforme Mittelwertsunterschied war jedoch nicht signifikant. Hierbei ist wiederum die geringe Teststärke von 10.1% zu kritisieren. Darüber hinaus ist zu kritisieren, dass der Fragebogen an dieser Stelle nicht eindeutig formuliert war (vgl. 4.1.4). Durch die bewusst offen gestaltete Frageform unterschieden sich die Angaben der Teilnehmenden stark. Eine andere Operationalisierung ist in zukünftigen Studien zu überdenken, so könnte unter Berücksichtigung des Aufwandes die tatsächliche Verhaltensänderung (statt der Vorsätze für die Verhaltensänderung) erfasst werden.

Hypothesenübergreifend sollen an dieser Stelle drei grundsätzliche Limitationen der Studie dargelegt werden: Erstens konnte trotz a priori festgestellter Stichprobengröße keine angemessen große Stichprobe erzielt werden, sondern nur 54.7% [52.3%] der a priori berechneten Teilnehmendenzahl. Dies ist im Wesentlichen durch die zeitlichen und finanziellen Rahmenbedingungen einer Masterarbeit bedingt. Für zukünftige Untersuchungen in diesem Forschungsbereich sollte mit den zu erwartenden Effekten angemessen großen Stichproben gearbeitet werden. Zweitens ist an dem Design zu kritisieren, dass die Beobachtenden beide Feedbackvarianten kannten und somit die Doppelblindheit nicht vollständig erfüllt war: Die Beobachtenden kannten zwar die Hypothesen nicht, da aber alle Psychologiestudierende waren, ist anzunehmen, dass sie Vermutungen hinsichtlich der Hypothesen entwickelt hatten. Dies stellt eine Einschränkung der internen Validität dar. Hinzu kommt, dass die Beobachtenden unterschiedlich häufig in den verschiedenen Bedingungen beobachtet haben und demnach Störeffekte nicht vollständig auszuschließen sind. Für zukünftige Untersuchungsdesigns ist eine stärkere Standardisierung des Feedbacks empfehlenswert, wenn auch die Umsetzung dessen schwierig erscheint, wenn man die externe Validität berücksichtigen möchte und deshalb mündliches Feedback (welches im Arbeitskontext häufig auftritt) zum

Untersuchungsgegenstand macht. Drittens war in diesem Design keine Manipulationskontrolle möglich. Idealerweise hätte kontrolliert werden müssen, ob die Beobachtenden während der Rückmeldung des Feedbacks sich an die zufällig zugeteilte Feedbackvariante halten und z. B. auch, ob sie in der einen Variante das Feedback freundlicher übermitteln als in der anderen (z. B. über Kodierung von Mimik). Dies war aufgrund der zeitlichen Rahmenbedingungen einer Masterarbeit nicht möglich, sollte aber für zukünftige Untersuchungen berücksichtigt werden.

4.3.2 Diskussion zu Selbstwertschätzung und Feedback

Neben den Hypothesen wurden in Studie 2 explorativ die Zusammenhänge zwischen Selbstwertschätzung und den verschiedenen Feedbackkonzepten untersucht. Die Vermutung, dass sich vor allem bei Personen mit niedriger Selbstwertschätzung Vorteile der Feedbackformulierung *beibehalten vs. verändern* zeigen, konnte so nicht bestätigt werden. Dennoch zeigen sich signifikante Unterschiede in der Anzahl der gefassten Vorsätze: Personen mit niedriger Selbstwertschätzung fassten im Mittel mehr Vorsätze als Personen mit hoher Selbstwertschätzung. Dies ist angesichts bisheriger Studien (Brockner, 1979; Brockner et al., 1983; Brockner et al., 1987; Brown & Marshall, 2001; Brown, 2010; Shrauger & Rosenberg, 1970; Shrauger & Sorman, 1977) zu negativen Konsequenzen von negativem Feedback überraschend (vgl. Kapitel 2.1.2.1), könnte jedoch mit der starken Verhaltensverankerung der Beobachtungsbögen in Zusammenhang stehen. Möglicherweise hilft die Verhaltensverankerung Personen mit niedrigem Selbstwert besonders dabei, konstruktiv mit Feedback umzugehen und Verhaltensvorsätze abzuleiten. Dennoch muss an dieser Stelle auch der Interaktionseffekt berücksichtigt werden: Personen mit hoher Selbstwertschätzung fassten in der BV-Bedingung mehr Vorsätze als diejenigen mit niedriger Selbstwertschätzung, in der GS-Bedingung wiederum fassten sie weniger Vorsätze. Dies widerspricht der Vermutung, dass die Formulierung von Feedback nach dem Konzept *beibehalten vs. verändern* vor allem bei Personen mit niedriger Selbstwertschätzung zu mehr Verhaltensintentionen führt – im Gegenteil: Die Formulierung *beibehalten vs. verändern* führt vor allem bei Personen mit hoher Selbstwertschätzung zu mehr Verhaltensintentionen. Grund dafür könnte sein, dass Personen mit hoher Selbstwertschätzung in der Formulierungsvariante Veränderungsvorschläge besser annehmen können als Feedback, das mit der Wertigkeit *schlecht* übermittelt wird und somit mehr Verhaltensintentionen generieren.

In der Konsequenz ist eine Berücksichtigung der Selbstwertschätzung im Kontext von Feedback für zukünftige Studien äußerst empfehlenswert, wenn auch auf Basis dieser Studie keine klare Aussage über den Zusammenhang getroffen werden kann. Analog zu Kapitel 4.3.1 sollte bei den zukünftigen Untersuchungen im Bereich Selbstwertschätzung und Feedback eine den zu erwartenden Effekten angemessene Stichprobengröße herangezogen werden.

5 Gesamtdiskussion

Wenn Feedback nach dem Formulierungskonzept *beibehalten vs. verändern* gegeben wird, nennen die Feedbackgebenden signifikant weniger eigenschaftsbezogene Inhalte als diejenigen, die nach dem Konzept *gut vs. schlecht* Feedback geben. Diejenigen, die Feedback nach dem Konzept *beibehalten vs. verändern* erhalten, weisen eine signifikant positivere Einstellung dem Feedback gegenüber auf. Personen mit niedriger Selbstwertschätzung fassen mehr Vorsätze als Personen mit hoher Selbstwertschätzung, wobei sich dieser Effekt zwischen den Feedbackkonzepten unterschiedet: Während Personen mit hoher Selbstwertschätzung nach dem Konzept *beibehalten vs. verändern* mehr Vorsätze fassen, fassen Personen mit niedriger Selbstwertschätzung nach dem Konzept *gut vs. schlecht* mehr Vorsätze.

Die Formulierung von Feedbacks mittels *beibehalten vs. verändern* zeigt sowohl auf der Seite der Feedbackgebenden als auch auf der Seite der Feedbacknehmenden positive Wirkung. Es zeichnet sich die Tendenz ab, dass das neue Formulierungskonzept einen fruchtbaren und vor allem praktischen Ansatz zur Bewältigung der paradoxen Effekte von Feedback sowie der Diskrepanz zwischen Feedbackkultur in Unternehmen und stark variierenden Effektstärken in wissenschaftlichen Studien darstellt.

Die vorliegenden Studien weisen aber sowohl methodische, als auch inhaltliche Ansatzpunkte zur Verbesserung auf, die in zukünftigen Untersuchungen berücksichtigt werden sollten.

Auf der methodischen Seite sind neben der bereits diskutierten Problematik der Stichprobengröße auch die verwendeten Kodierverfahren zu nennen, die zwar eine Näherung an die Thematik, aber wohl keine optimale Lösung darstellen. Darüber hinaus weist das Studiendesign in Studie 2 diverse Schwachstellen auf – u. a. wie bereits diskutiert die Verletzung der Doppelblind-Bedingung der Beobachtenden. Hier könnte man ein alternatives Studiendesign entwickeln. Kritisch ist in dem Zusammenhang jedoch immer, dass zunächst eine Leistung erbracht werden muss, die im Anschluss rückgemeldet werden kann. Durch diese Mehrstufigkeit im Verfahren ergeben sich zwangsläufig Fehlerquellen, deren Kontrolle sich schwierig gestaltet (vgl. Price et al. (2010)).

Inhaltlich bieten sich verschiedene Weiterentwicklungen an. Einerseits könnte die Theorie stärker in das gut abgesicherte Modell der Forschungsgruppe um Carol Dweck zum stati-

schen vs. dynamischen Denkkonzept eingebettet werden. Anknüpfungspunkte zur Integration liefern u. a. Dweck et al. (1995a) mit den Überlegungen zum Zusammenhang von statischem und dynamischen Denkkonzept mit Persönlichkeit, Motivation und sozialer Wahrnehmung. Die Formulierung von Feedback könnte in diesen Zusammenhängen eine moderierende oder mediierende Rolle spielen. Andererseits wären auch weiterführende Untersuchungen von Interesse, die die Formulierung von Feedback im Kontext hoher vs. niedriger Selbstwertschätzung untersuchen. Möglicherweise könnte so ein Teil der Varianz in Effektrichtungen und -stärken aufgeklärt und damit Klarheit bezüglich der Wirksamkeit von Feedback geschaffen werden.

Die Formulierung von Feedback als *beibehalten* vs. *verändern* ist ein neues Konzept, das durch weitere Studien stärker im Theoriekontext verankert sowie weiterführend empirisch überprüft werden sollte und so die Feedbackkultur voranbringen kann.

Literaturverzeichnis

- Aguinis, H., Gottfredson, R. K. & Joo, H. (2012). Delivering effective performance feedback. The strengths-based approach. *Business Horizons*, 55 (2), 105–111.
<https://doi.org/10.1016/j.bushor.2011.10.004>
- Arbeitskreis Assessment Center e.V. (2016). *AC-Standards. Standards der Assessment Center Methode* (3. vollständig überarbeitete Fassung). Verfügbar unter <https://www.arbeitskreis-ac.de/images/attachments/AkAC-Standards-2016.pdf>
- Ashford, S., Blatt, R. & VandeWalle, D. (2003). Reflections on the Looking Glass. A Review of Research on Feedback-Seeking Behavior in Organizations. *Journal of Management*, 29 (6), 773–799. [https://doi.org/10.1016/S0149-2063\(03\)00079-5](https://doi.org/10.1016/S0149-2063(03)00079-5)
- Assor, A., Roth, G. & Deci, E. L. (2004). The Emotional Costs of Parents' Conditional Regard. A Self-Determination Theory Analysis. *Journal of Personality*, 72 (1), 47–88.
<https://doi.org/10.1111/j.0022-3506.2004.00256.x>
- Besieux, T. (2017). Why I hate feedback. Anchoring effective feedback within organizations. *Business Horizons*, 60 (4), 435–439. <https://doi.org/10.1016/j.bushor.2017.03.001>
- Blackwell, L. S., Trzesniewski, K. H. & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition. A longitudinal study and an intervention. *Child development*, 78 (1), 246–263. <https://doi.org/10.1111/j.1467-8624.2007.00995.x>
- Boneau, C. A. (1960). The effects of violations of assumptions underlying the t test. *Psychological Bulletin*, 57 (1), 49–64. <https://doi.org/10.1037/h0041412>
- Bortz, J. & Schuster, C. (2016). *Statistik für Human- und Sozialwissenschaftler. Extras online* (Springer-Lehrbuch, Limitierte Sonderausgabe, 7., vollständig überarbeitete und erweiterte Auflage). Berlin: Springer.
- Bouskila-Yam, O. & Kluger, A. N. (2011). Strength-based performance appraisal and goal setting. *Human Resource Management Review*, 21 (2), 137–147.
<https://doi.org/10.1016/j.hrmr.2010.09.001>
- Brinko, K. T. (1993). The Practice of Giving Feedback to Improve Teaching. *The Journal of Higher Education*, 64 (5), 574–593.
<https://doi.org/10.1080/00221546.1993.11778449>

- Brockner, J., Gardner, M., Bierman, J., Mahan, T., Thomas, B., Weiss, W. et al. (1983). The roles of self-esteem and self-consciousness in the Wortman-Brehm model of reactance and learned helplessness. *Journal of Personality and Social Psychology*, 45 (1), 199–209.
- Brockner, J. (1979). The effects of self-esteem, success-failure, and self-consciousness on task performance. *Journal of Personality and Social Psychology*, 37 (10), 1732–1741. <https://doi.org/10.1037/0022-3514.37.10.1732>
- Brockner, J., Derr, W. R. & Laing, W. N. (1987). Self-esteem and reactions to negative feedback. Toward greater generalizability. *Journal of Research in Personality*, 21 (3), 318–333. [https://doi.org/10.1016/0092-6566\(87\)90014-6](https://doi.org/10.1016/0092-6566(87)90014-6)
- Brown, J. D. (2010). High self-esteem buffers negative feedback: Once more with feeling. *Cognition & Emotion*, 24 (8), 1389–1404. <https://doi.org/10.1080/02699930903504405>
- Brown, J. D. & Marshall, M. A. (2001). Self-Esteem and Emotion: Some Thoughts about Feelings. *Personality & social psychology bulletin*, 27 (5), 575–584. <https://doi.org/10.1177/0146167201275006>
- Brummelman, E., Thomaes, S., Orobio de Castro, B., Overbeek, G. & Bushman, B. J. (2014). "That's not just beautiful--that's incredibly beautiful!". The adverse impact of inflated praise on children with low self-esteem. *Psychological science*, 25 (3), 728–735. <https://doi.org/10.1177/0956797613514251>
- Brummelman, E., Thomaes, S., Overbeek, G., Orobio de Castro, B., van den Hout, M. A. & Bushman, B. J. (2014). On feeding those hungry for praise. Person praise backfires in children with low self-esteem. *Journal of experimental psychology. General*, 143 (1), 9–14. <https://doi.org/10.1037/a0031917>
- Burkert, S. (2017). Kybernetische Feedbackschleifenmodelle. In M. A. Wirtz (Hrsg.), *Dorsch - Lexikon der Psychologie* (18., überarbeitete Auflage). Bern: Hogrefe. Zugriff am 18.07.2018. Verfügbar unter <https://m.portal.hogrefe.com/dorsch/kybernetische-feedbackschleifenmodelle/>
- Busemann, A. (1925). Die Sprache der Jugend als Ausdruck der Entwicklungsrhythmik. Sprachstatistische Untersuchungen. In C. Bühler (Hrsg.) Quellen und Studien zur Jugendkunde, 1–108 [Themenheft]. Jena: Gustav Fischer.
- Canavan, C., Holtman, M. C., Richmond, M. & Katsufakis, P. J. (2010). The quality of written comments on professional behaviors in a developmental multisource feedback program. *Academic medicine : journal of the Association of American Medical Colleges*, 85 (10 Suppl), S106-9. <https://doi.org/10.1097/ACM.0b013e3181ed4cdb>

- Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31* (2), 219–233. <https://doi.org/10.1080/03075070600572132>
- Carver, C. S. & Scheier, M. F. (1981). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology, 17* (6), 545–568. [https://doi.org/10.1016/0022-1031\(81\)90039-1](https://doi.org/10.1016/0022-1031(81)90039-1)
- Cimpian, A., Arce, H.-M. C., Markman, E. M. & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological science, 18* (4), 314–316. <https://doi.org/10.1111/j.1467-9280.2007.01896.x>
- Cimpian, A. & Markman, E. M. (2011). The generic/nongeneric distinction influences how children interpret new information about social others. *Child development, 82* (2), 471–492. <https://doi.org/10.1111/j.1467-8624.2010.01525.x>
- Cohen, J. (1983). The Cost of Dichotomization. *Applied Psychological Measurement, 7* (3), 249–253. <https://doi.org/10.1177/014662168300700301>
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction. A meta-analysis of findings. *Research in Higher Education, 13* (4), 321–341. <https://doi.org/10.1007/BF00976252>
- Collani, G. von & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift für Differentielle und Diagnostische Psychologie, 24* (1), 3–7. <https://doi.org/10.1024//0170-1789.24.1.3>
- Corpus, J. H., Ogle, C. M. & Love-Geiger, K. E. (2006). The Effects of Social-Comparison Versus Mastery Praise on Children's Intrinsic Motivation. *Motivation and Emotion, 30* (4), 333–343. <https://doi.org/10.1007/s11031-006-9039-4>
- Dimotakis, N., Mitchell, D. & Maurer, T. (2017). Positive and negative assessment center feedback in relation to development self-efficacy, feedback seeking, and promotion. *The Journal of applied psychology, 102* (11), 1514–1527. <https://doi.org/10.1037/apl0000228>
- Donnelly, P. & Kirk, P. (2015). How to... Give effective feedback. *Education for Primary Care, 21* (4), 267–269. <https://doi.org/10.1080/14739879.2010.11493919>
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (Springer-Lehrbuch, 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Berlin: Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41* (10), 1040–1048. <https://doi.org/10.1037//0003-066X.41.10.1040>

- Dweck, C. S. (2002). The Development of Ability Conceptions. In A. Wigfield & J. Eccles (Hrsg.), *Development of achievement motivation. A volume in Educational Psychology* (S. 57–88). <https://doi.org/10.1016/B978-012750053-9/50005-X>
- Dweck, C. S., Chiu, C.-y. & Hong, Y.-y. (1995a). Implicit Theories and Their Role in Judgments and Reactions. A Word From Two Perspectives. *Psychological Inquiry*, 6 (4), 267–285. https://doi.org/10.1207/s15327965pli0604_1
- Dweck, C. S., Chiu, C.-y. & Hong, Y.-y. (1995b). Implicit Theories. Elaboration and Extension of the Model. *Psychological Inquiry*, 6 (4), 322–333. https://doi.org/10.1207/s15327965pli0604_12
- Dweck, C. S. & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95 (2), 256–273. <https://doi.org/10.1037/0033-295X.95.2.256>
- Echterhoff, G. (2012). Shared-Reality Theory. In A. W. Kruglanski, E. T. Higgins & P. A. M. van Lange (Eds.), *Theories of social psychology* (pp. 180–199). Los Angeles, Calif.: Sage. <https://doi.org/10.4135/9781446249222.n35>
- Echterhoff, G., Higgins, E. T. & Levine, J. M. (2009). Shared Reality: Experiencing Commonality With Others' Inner States About the World. *Perspectives on psychological science : a journal of the Association for Psychological Science*, 4 (5), 496–521. <https://doi.org/10.1111/j.1745-6924.2009.01161.x>
- Elliott, E. S. & Dweck, C. S. (1988). Goals. An approach to motivation and achievement. *Journal of Personality and Social Psychology*, 54 (1), 5–12. <https://doi.org/10.1037/0022-3514.54.1.5>
- Evans, C. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, 83 (1), 70–120. <https://doi.org/10.3102/0034654312474350>
- Fedor, D. B., Eder, R. W. & Buckley, M.R. (1989). The contributory effects of supervisor intentions on subordinate feedback responses. *Organizational Behavior and Human Decision Processes*, 44 (3), 396–414. [https://doi.org/10.1016/0749-5978\(89\)90016-2](https://doi.org/10.1016/0749-5978(89)90016-2)
- Fedor, D. B., Rensvold, R. B. & Adams, S. M. (1992). An investigation of factors expected to affect feedback seeking. A longitudinal field study. *Personnel Psychology*, 45 (4), 779–802. <https://doi.org/10.1111/j.1744-6570.1992.tb00968.x>
- Field, A., Miles, J. & Field, Z. (2012). *Discovering statistics using R*. Los Angeles: Sage.
- Gabelica, C., van den Bossche, P., Segers, M. & Gijssels, W. (2012). Feedback, a powerful lever in teams. A review. *Educational Research Review*, 7 (2), 123–144. <https://doi.org/10.1016/j.edurev.2011.11.003>

- Gibbs, G. & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education* (1), 3–31.
- Gillie, P. J. (1957). A simplified formula for measuring abstraction in writing. *Journal of Applied Psychology*, 41 (4), 214–217. <https://doi.org/10.1037/h0045036>
- Glass, G. V., Peckham, P. D. & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, 42 (3), 237–288. <https://doi.org/10.3102/00346543042003237>
- Gollwitzer, P. M. & Sheeran, P. (2006). Implementation Intentions and Goal Achievement. A Meta-analysis of Effects and Processes. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Advances in Experimental Social Psychology, vol. 38, 1st ed., Bd. 38, pp. 69–119). s.l.: Elsevier textbooks. [https://doi.org/10.1016/S0065-2601\(06\)38002-1](https://doi.org/10.1016/S0065-2601(06)38002-1)
- Greve, W., Wentura, D., Gräser, H. & Schmitz, U. (1997). *Wissenschaftliche Beobachtung. Eine Einführung* [2. Aufl.]. Weinheim: Beltz.
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S. & Levine, S. C. (2013). Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later. *Child development*, 84 (5), 1526–1541. <https://doi.org/10.1111/cdev.12064>
- Günther, U. L. & Groeben, N. (1978). Abstrakheitssuffix-Verfahren: Vorschlag einer objektiven ökonomischen Messung der Abstraktheit/Konkretheit von Texten. *Zeitschrift für experimentelle und angewandte Psychologie*, 25 (1), 55–74.
- Hamilton, B. L. (1977). An Empirical Investigation of the Effects of Heterogeneous Regression Slopes in Analysis of Covariance. *Educational and Psychological Measurement*, 37 (3), 701–712. <https://doi.org/10.1177/001316447703700313>
- Harwell, M. R., Rubinstein, E. N., Hayes, W. S. & Olds, C. C. (2016). Summarizing Monte Carlo Results in Methodological Research. The One- and Two-Factor Fixed Effects ANOVA Cases. *Journal of Educational Statistics*, 17 (4), 315–339. <https://doi.org/10.3102/10769986017004315>
- Hattie, J. (1999, August). *Influences on student learning*. Inaugural Lecture: Professor of Education. Zugriff am 10.02.2018.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77 (1), 81–112. <https://doi.org/10.3102/003465430298487>

- Hong, Y.-y., Chiu, C.-y., Dweck, C. S., Lin, D. M.-S. & Wan, W. (1999). Implicit theories, attributions, and coping. A meaning system approach. *Journal of Personality and Social Psychology*, 77 (3), 588–599. <https://doi.org/10.1037/0022-3514.77.3.588>
- Ilgen, D. R., Fisher, C. D. & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *The Journal of applied psychology*, 64 (4), 349–371. <https://doi.org/10.1037/0021-9010.64.4.349>
- Irmen, L. & Linner, U. (2005). Die Repräsentation generisch maskuliner Personenbezeichnungen. *Zeitschrift für Psychologie / Journal of Psychology*, 213 (3), 167–175. <https://doi.org/10.1026/0044-3409.213.3.167>
- Kamins, M. L. & Dweck, C. S. (1999). Person versus process praise and criticism. Implications for contingent self-worth and coping. *Developmental psychology*, 35 (3), 835–847. <https://doi.org/10.1037/0012-1649.35.3.835>
- Kluger, A. N. & DeNisi, A. (1996). The effects of feedback interventions on performance. A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119 (2), 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159–174.
- Lix, L. M., Keselman, J. C. & Keselman, H. J. (1996). Consequences of Assumption Violations Revisited. A Quantitative Review of Alternatives to the One-Way Analysis of Variance "F" Test. *Review of Educational Research*, 66 (4), 579. <https://doi.org/10.2307/1170654>
- McKimm, J. (2009). Giving effective feedback. *British journal of hospital medicine (London, England : 2005)*, 70 (3), 158–161. <https://doi.org/10.12968/hmed.2009.70.3.40570>
- Miller, G. A. & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110 (1), 40–48. <https://doi.org/10.1037/0021-843X.110.1.40>
- Mueller, C. M. & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75 (1), 33–52.
- Overall, J. E. & Woodward, J. A. (1977a). Common Misconceptions Concerning The Analysis Of Covariance. *Multivariate behavioral research*, 12 (2), 171–186. https://doi.org/10.1207/s15327906mbr1202_5
- Overall, J. E. & Woodward, J. A. (1977b). Nonrandom assignment and the analysis of covariance. *Psychological Bulletin*, 84 (3), 588–594. <https://doi.org/10.1037/0033-2909.84.3.588>

- Pomerantz, E. M. & Kempner, S. G. (2013). Mothers' daily person and process praise. Implications for children's theory of intelligence and motivation. *Developmental psychology*, 49 (11), 2040–2046. <https://doi.org/10.1037/a0031840>
- Price, M., Handley, K., Millar, J. & O'Donovan, B. (2010). Feedback. All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35 (3), 277–289. <https://doi.org/10.1080/02602930903541007>
- Prins, F. J., Sluijsmans, D. M. A. & Kirschner, P. A. (2006). Feedback for general practitioners in training. Quality, styles, and preferences. *Advances in health sciences education : theory and practice*, 11 (3), 289–303. <https://doi.org/10.1007/s10459-005-3250-z>
- Institut für Deutsche Sprache. (2018). *Geschlechtergerechte Schreibung: Herausforderung noch ohne Lösung*. Mannheim. Verfügbar unter http://www.rechtschreib-rat.com/DOX/rfdr_PM_2018-06-08_Geschlechtergerechte_Schreibung.pdf
- Rattan, A., Good, C. & Dweck, C. S. (2012). “It's ok — Not everyone can be good at math”. Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, 48 (3), 731–737. <https://doi.org/10.1016/j.jesp.2011.12.012>
- Rhodewalt, F. (1994). Conceptions of Ability, Achievement Goals, and Individual Differences in Self-Handicapping Behavior. On the Application of Implicit Theories. *Journal of Personality*, 62 (1), 67–85. <https://doi.org/10.1111/j.1467-6494.1994.tb00795.x>
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology*, 17 (4), 688–690. <https://doi.org/10.1093/beheco/ark016>
- Shrauger, J. S. (1972). Self-esteem and reactions to being observed by others. *Journal of Personality and Social Psychology*, 23 (2), 192–200. <https://doi.org/10.1037/h0033046>
- Shrauger, J. S. & Rosenberg, S. E. (1970). Self-esteem and the effects of success and failure feedback on performance. *Journal of Personality*, 38 (3), 404–417. <https://doi.org/10.1111/j.1467-6494.1970.tb00018.x>
- Shrauger, J. S. & Sorman, P. B. (1977). Self-evaluations, initial success and failure, and improvement as determinants of persistence. *Journal of Consulting and Clinical Psychology*, 45 (5), 784–795. <https://doi.org/10.1037/0022-006X.45.5.784>
- Slowiak, J. M. & Lakowske, A. M. (2017). The influence of feedback statement sequence and goals on task performance. *Behavior Analysis: Research and Practice*, 17 (4), 357–380. <https://doi.org/10.1037/bar0000084>

- Sluijsmans, D. M. A., Brand-Gruwel, S. & van Merriënboer, J. J. G. (2002). Peer Assessment Training in Teacher Education. Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, 27 (5), 443–454.
<https://doi.org/10.1080/0260293022000009311>
- Steelman, L. A. & Rutkowski, K. A. (2004). Moderators of employee reactions to negative feedback. *Journal of Managerial Psychology*, 19 (1), 6–18.
<https://doi.org/10.1108/02683940410520637>
- Van der Leeuw, R. M., Schipper, M. P., Heineman, M. J. & Lombarts, K. M. J. M. H. (2016). Residents' narrative feedback on teaching performance of clinical teachers. Analysis of the content and phrasing of suggestions for improvement. *Postgraduate medical journal*, 92 (1085), 145–151. <https://doi.org/10.1136/postgradmedj-2014-133214>
- Vester, H.-G. (2009). *Kompodium der Soziologie I: Grundbegriffe* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden.
<https://doi.org/10.1007/978-3-531-91345-2>
- Villiers, R. de. (2013). 7 Principles of highly effective managerial feedback. Theory and practice in managerial development interventions. *The International Journal of Management Education*, 11 (2), 66–74. <https://doi.org/10.1016/j.ijme.2013.01.002>
- Welch, B. L. (1938). The Significance of the Difference Between Two Means when the Population Variances are Unequal. *Biometrika*, 29 (3/4), 350.
<https://doi.org/10.2307/2332010>

Anhang

Anhang A: Einladungsmail zur Teilnahme aus Studie 1.....	69
Anhang B: Fragebogen der BV-Bedingung aus Studie 1.....	70
Anhang C: Ausschnitt des Fragebogens der GS-Bedingung aus Studie 1.....	75
Anhang D: Ratingregeln aus Studie 1.....	77
Anhang E: Beobachtungsbögen aus Studie 2.....	79
Anhang F: Beobachtungsinstruktionen aus Studie 2.....	81
Anhang G: Feedbackinstruktionen aus Studie 2.....	83
Anhang H: Materialienliste aus Studie 2.....	85
Anhang I: Teilnehmendeninstruktionen aus Studie 2.....	86
Anhang J: Höchster Turm mit 106 cm aus Studie 2.....	87
Anhang K: Teilnehmendenwerbung aus Studie 2.....	88
Anhang L: Fragebogen 1 aus Studie 2.....	89
Anhang M: Fragebogen 2 aus Studie 2.....	90
Anhang N: Fragebogen 3 aus Studie 2.....	91
Anhang O: Kodierregeln aus Studie 2.....	93

Anhang A: Einladungsmail zur Teilnahme aus Studie 1

Liebe Kursteilnehmer/-innen,

unter folgendem Link finden Sie einen Feedback-Fragebogen zu dieser Veranstaltung: <https://www.soscisurvey.de/kursfeedback/?q=DidaktikSachunterricht>

Für eine stetige Weiterentwicklung von Form und Inhalten ist Ihre Rückmeldung sehr wichtig für mich.

Bitte füllen Sie - in Ihrem eigenen Interesse und dem Interesse nachfolgender Studierender - den Fragebogen nach bestem Wissen und Gewissen bis **spätestens Mittwoch, den 20.06.2018**, aus.

Der Fragebogen wurde entwickelt von Ihrer Kommilitonin Anna-Katharina Stöcker, die am Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik der Uni Bamberg ihre Masterarbeit schreibt. Mit der Teilnahme an diesem Feedback unterstützen Sie nicht nur die Weiterentwicklung dieses Kurses, sondern auch die Masterarbeit von Frau Stöcker.

Vielen herzlichen Dank vorab,

mit freundlichen Grüßen

Anhang B: Fragebogen der BV-Bedingung aus Studie 1

Otto-Friedrich-Universität Bamberg



Feedback zur Veranstaltung „ISDL2 – Optimierung IT-lastiger Geschäftsprozesse“

am 27.06.2018

1. Einverständniserklärung

Da die Angaben in diesem Feedback auch im Rahmen einer Studie zum Thema "Feedback" am Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik, Personal- und Sozialpsychologie ausgewertet werden, benötigen wir neben einigen Einverständniserklärungen auch einzelne allgemeine Angaben von dir.

Die Angaben zur Person werden selbstverständlich nicht an Lea Müller rückgemeldet, lediglich das Feedback wird anonymisiert und als Zusammenfassung aller Teilnehmendenangaben zurückgemeldet.

Die Teilnahme an dieser Studie ist freiwillig. Ein Abbruch der Studie ist jederzeit und ohne negative Konsequenzen für mich möglich. Ich habe jederzeit mit Wirkung für die Zukunft das Recht, die Einverständniserklärung zu widerrufen. Ein Widerruf ist möglich mittels einer E-Mail an Frau Anna-Katharina Stöcker (anna-katharina-christa.stoecker@stud.uni-bamberg.de). Im Fall des Widerrufs werden mit dem Zugang meiner Widerrufserklärung meine Daten am Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik gelöscht.

- Ich erkläre mich **einverstanden**, an der Studie teilzunehmen.
- Ich erkläre mich **nicht einverstanden**, an der Studie teilzunehmen.

Die Ergebnisse und Daten werden im Rahmen der Masterarbeit von Frau Anna-Katharina Stöcker ausgewertet. Ihre Kontaktdaten findest du auf der letzten Seite des Fragebogens.

- Ich erkläre mich **einverstanden**, dass meine vollständig anonymisierten Daten im Rahmen der Masterarbeit von Frau Anna-Katharina Stöcker ausgewertet werden.
- Ich erkläre mich **nicht einverstanden**, dass meine vollständig anonymisierten Daten im Rahmen der Masterarbeit von Frau Anna-Katharina Stöcker ausgewertet werden.

Die Ergebnisse und Daten dieser Studie werden als wissenschaftliche Publikation veröffentlicht. Dies geschieht in anonymisierter Form, d. h. ohne, dass die Daten einer spezifischen Person zugeordnet werden können.

- Ich erkläre mich **einverstanden**, dass meine vollständig anonymisierten Daten als offene Daten im Internet in einem wissenschaftlichen Datenarchiv zugänglich gemacht werden. Damit folgt diese Studie den Empfehlungen der Deutschen Forschungsgemeinschaft (DFG) und der Deutschen Gesellschaft für Psychologie (DGPs) zur Qualitätssicherung in der Forschung.
- Ich erkläre mich **nicht einverstanden**, dass meine vollständig anonymisierten Daten als offene Daten im Internet in einem wissenschaftlichen Datenarchiv zugänglich gemacht werden.

2. Allgemeine Angaben

Geschlecht: männlich
 weiblich
 ich möchte keine Angabe machen

Studiengang: _____

Grund des Besuchs der Veranstaltung: Pflichtveranstaltung
 freiwilliger Besuch
 Sonstiges: _____

Bitte gib auf einer Skala von 1 bis 7 an, wie sehr du der folgenden Aussage zustimmst:

	1	2	3	4	5	6	7
	<small>überhaupt nicht</small>						<small>voll und ganz</small>
„Ich gebe generell gerne Feedback.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. Feedback

- Bitte gib auf den folgenden Seiten dein Feedback zur Veranstaltung „ISDL2 – Optimierung IT-lastiger Geschäftsprozesse“ ab.
- Bitte fülle den Erhebungsbogen nach bestem Wissen aus, **Nachfragen sind nicht möglich**.
- Deine Angaben werden anschließend systematisiert an Lea Müller rückgemeldet.

Notiere an den jeweiligen Stichpunkten, was deiner Meinung nach **beibehalten werden sollte**.

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Ich finde, es sollte beibehalten werden, dass _____

Notiere an den jeweiligen Stichpunkten, was deiner Meinung nach **verändert werden sollte (wovon mehr? wovon weniger? was anders?)**.

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

Ich finde, es sollte verändert werden, dass _____

3. Abschließende Fragen

Bitte gib auf einer Skala von 1 bis 7 an, wie sehr du der folgenden Aussage zustimmst:

	1	2	3	4	5	6	7
	<small>überhaupt nicht</small>						<small>voll und ganz</small>
„Ich habe dieses Mal gerne Feedback gegeben.“	0	0	0	0	0	0	0

Mein Grund für diese Angabe: _____

Platz für sonstige Kommentare zur Erhebung:

Kontaktdaten:

Lehrstuhl für Persönlichkeitspsychologie und Psychologische Diagnostik

Otto-Friedrich-Universität Bamberg

Studienverantwortliche: *Anna-Katharina Stöcker*

Markusplatz 3

96047 Bamberg

E-Mail-Adresse: *anna-katharina-christa.stoecker@stud.uni-bamberg.de*

Anhang C: Ausschnitt des Fragebogens der GS-Bedingung aus Studie 1

Notiere an den jeweiligen Stichpunkten, was deiner Meinung nach **gut ist**.

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Ich finde gut, dass _____

Notiere an den jeweiligen Stichpunkten, was deiner Meinung nach **schlecht ist**.

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Ich finde schlecht, dass _____

Anhang D: Ratingregeln aus Studie 1

1. Einhalten des grammatikalischen Schemas

- Trage in das Feld gr die Bewertung als grammatikalisch korrekt (=0) oder falsch (=1) ein bzw. „2“ falls die Vpn z.T. korrekte, z.T. falsche grammatikalische Aussagen trifft
- Korrekt ist der Satz dann, wenn er eine korrekte Ergänzung des Satzes „Ich finde gut, dass...“ ist.

2. Auszählen von Wortarten

- Substantive
 - Trage in das Feld ko_nu1 die Anzahl der Abstraktheits-Substantive ein. Es zählen alle deutschen Substantive dazu, die folgende Endungen aufweisen:
 - heit, z. B. Gewohnheit
 - ie, z. B. Manie (aber nur, wenn es wie ein langes „i“ gesprochen wird; nicht, wenn es wie „je“ gesprochen wird, z. B. Kastanie)
 - ik, z. B. Rhythmik
 - ion, z. B. Konzentration
 - ismus, z. B. Protestantismus
 - ität, z. B. Identität
 - keit, z. B. Häufigkeit
 - nz, z. B. Potenz
 - tur, z. B. Struktur
 - ung, z. B. Hoffnung
 - sowie alle Substantive der englischen Sprache, die folgende Endungen aufweisen:
 - ness
 - ment
 - ship
 - dom
 - nce
 - ion
 - y (allerdings nicht, wenn es ein Diminutiv ist)
 - Trage in das Feld su_nu1 die Anzahl der Substantive insgesamt ein.
 - Beispiele und Sonderfälle:
 - Namen zählen weder als Substantiv, noch werden sie hinsichtlich der Abstraktheit bewertet
 - Case Study: zwei Substantive, davon ein Abstraktheitssubstantiv
 - AC-Training: ein Substantiv
 - Ziel- und Ergebnisorientierung: zwei Substantive, zwei Abstraktheitssubstantive
 - AC = Assessment Center: zwei Substantive, ein Abstraktheitssubstantiv
- Verben und Adjektive
 - Trage in das Feld ak_nu1 die Anzahl aktionaler Aussagen ein. Dazu zählen:
 - Alle Aussagen mit finitem Verb und im Infinitiv mit „um zu“ und „zu“, wenn sie nicht bloße Ortsbestimmungen, Besitzanzeigen, Existenzialaussagen oder Ganzes-Teil-Aussagen darstellen, z. B. „er geht“ (a), „er geht (a), um etwas zu holen (a)“, „er versucht (a) zu schlafen (a)“
 - Infinitive mit den Hilfszeitwörtern „können, wollen, mögen, sollen, dürfen“ werden ebenso behandelt, z. B. „er kann gehen (a)“
 - Trage in das Feld qu_nu1 die Anzahl qualitativer Aussagen ein. Dazu zählen
 - Alle Aussagen mit der Kopula (Hilfszeitwort „sein“), auch wenn das Hilfszeitwort im Infinitiv steht, z. B. „er ist müde (q)“, „was ist er (q)“, „es soll so sein (q)“

- Alle Appositionen (z. B. Karl *der Große*, die *behandelnde* Ärztin) und attributiven Adjektive, Ordinalzahlen, Partizipien und Substantive, z. B. „der kranke (q) Mann“, „Berlin, die Hauptstadt (q) des Reiches, ist schön (q)“, „der erste (q) Schuss“, „der reizende (q) Garten“
- Alle adverbiellen Adjektive
- Beachte:
 - Die Hilfszeitwörter „können, wollen mögen usw.“ in Verbindung mit „sein“ gelten an sich als aktionale Aussagen, z. B. „er kann (a) es sein (q)“
 - Die Hilfszeitwörter sowie das Hilfszeitwort „haben“ werden, wenn sie mit Zeitwörtern verbunden sind, nicht gezählt, z. B. „er kann lesen“ (=1a), „er kann lesen und schreiben (=2a)“, „er kann!“ (=1a)
 - Doppelzählung ist möglich, z. B. „er kann lesen und schreiben (=2a)“ und „es viel zu viel und unübersichtlich ist“ (=2q)

Anhang E: Beobachtungsbögen aus Studie 2

Teilnehmenden-Code: _____ Beobachtenden-Code: _____ Datum: _____ Uhrzeit: _____

Was sollte der*die Teilnehmende im Hinblick auf...	beibehalten:	verändern (mehr/ weniger/ anders):
Kooperation		
nonverbale Kommunikation		
Haltung		
Armhaltung (unverschränkt)		
Blickrichtung (zugewandt)		
Blickkontakt (aktiv suchend, haltend)		
Gestik (die Sprache unterstützend, angenehm)		
verbale Kommunikation		
Formulierung (präzise, angemessen)		
Verständlichkeit		
Lautstärke (angenehm)		
Artikulation (klar, deutlich)		
Zusammenarbeit		
Hilfsbereitschaft (reicht Gegenstände, unterstützt beim Bauen)		
Aufmerksamkeit (verweist auf genannte Punkte)		
Gesprächsverhalten (lässt andere ausreden)		
Wertschätzung (lobt Vorschläge/Einwände anderer)		
Ziel- und Ergebnisorientierung		
strukturiertes Arbeiten		
Strukturierung der Arbeitsabläufe (Aufgabenverteilung, Aufgabenpriorisierung, Teilschritte bis zum Ziel)		
Blick für's Große und Ganze (verweist auf Rahmenbedingungen, d.h. Zeit/ Material/ exakte Aufgabenstellung)		
strategische Handlungsweisen		
Vorantreiben des Prozesses (Zusammenfassung des Arbeitsstands und weiterführende Fragen)		
Initiierung von Verbesserungsprozessen zum richtigen Zeitpunkt		
Abschätzen der Auswirkungen von Handlungsweisen vor Durchführung		
effizientes Arbeiten		
zügige Arbeitsweise (ruhiges Abarbeiten der Aufgabenkomponenten, keine hektischen Armbewegungen)		

Teilnehmenden-Code: _____ Beobachtenden-Code: _____ Datum: _____ Uhrzeit: _____

Was hat der*die Teilnehmende im Hinblick auf...	gut gemacht:	schlecht gemacht:
Kooperation		
nonverbale Kommunikation		
Haltung Armhaltung (unverschränkt) Blickrichtung (zugewandt) Blickkontakt (aktiv suchend, haltend)		
Gestik (die Sprache unterstützend, angenehm)		
verbale Kommunikation		
Formulierung (präzise, angemessen)		
Verständlichkeit Lautstärke (angenehm) Artikulation (klar, deutlich)		
Zusammenarbeit		
Hilfsbereitschaft (reicht Gegenstände, unterstützt beim Bauen)		
Aufmerksamkeit (verweist auf genannte Punkte)		
Gesprächsverhalten (lässt andere ausreden)		
Wertschätzung (lobt Vorschläge/Einwände anderer)		
Ziel- und Ergebnisorientierung		
strukturiertes Arbeiten		
Strukturierung der Arbeitsabläufe (Aufgabenverteilung, Aufgabenpriorisierung, Teilschritte bis zum Ziel)		
Blick für's Große und Ganze (verweist auf Rahmenbedingungen, d.h. Zeit/ Material/ exakte Aufgabenstellung)		
strategische Handlungsweisen		
Vorantreiben des Prozesses (Zusammenfassung des Arbeitsstands und weiterführende Fragen)		
Initiierung von Verbesserungsprozessen zum richtigen Zeitpunkt		
Abschätzen der Auswirkungen von Handlungsweisen vor Durchführung		
effizientes Arbeiten		
zügige Arbeitsweise (ruhiges Abarbeiten der Aufgabenkomponenten, keine hektischen Armbewegungen)		

Anhang F: Beobachtungsinstruktionen aus Studie 2

Beobachtungsleitfaden

- Kopfzeile ausfüllen bitte nicht vergessen. Deinen Beobachtungscode erstellst du bitte nach folgendem Schema:

▪ Vorname der Mutter:	Martina	→ MA
▪ Geburtsort:	Bamberg	→ BA
▪ Geburtstag:	25.03.1994	→ 25
Das ergibt folgenden Code: MA BA 25		

- Nicht jedes Verhalten zeigt sich bei jedem*r – Verhalten, das nicht gezeigt wird, kann auch nicht beobachtet werden.
- Notiere im Beobachtungsbogen, welches Verhalten gezeigt wird. Halte dich an die vorgegebenen Verhaltensanker.

Was sollte der*die Teilnehmende im Hinblick auf...	beibehalten:	verändern (mehr/ weniger/ anders):
Kooperation		
nonverbale Kommunikation		
Haltung		
Armhaltung (unverschränkt)	<i>dem Gegenüber zugewandter</i>	<i>offenere Armhaltung, auch dem</i>
Blickrichtung (zugewandt)	<i>Blick, Blickkontakt haltend</i>	<i>Sitznachbarn den Blick</i>
Blickkontakt (aktiv suchend, haltend)		<i>zuwenden, Blickkontakt aktiv</i>
		<i>suchen</i>

- Wenn du angibst, dass der*die Teilnehmende sein*ihr Verhalten verändern sollte, konkretisiere dies bitte:
 - mehr: das Verhalten sollte häufiger oder stärker gezeigt werden, z.B.
 - Armhaltung: mehr Bewegung
 - Lautstärke: lauter
 - weniger: das Verhalten sollte seltener oder schwächer gezeigt werden, z.B.
 - Hilfsbereitschaft: unterstützt so sehr, dass er eigene Aufgaben nicht bewältigen kann → weniger helfen
 - Aufmerksamkeit: reagiert sehr aufmerksam, wirkt dadurch aber überangespannt
 - Gesprächsverhalten: lässt andere so sehr ausreden, dass er selbst nicht zu Wort kommt
 - anders: die Vpn sollte anderes Verhalten zeigen
 - Gestik: zeigt lebhaftes Gestik, die die Sprache aber nicht so gut unterstützt → präzisere Abstimmung zwischen Tonspur und Gestik nötig
 - Initiierung von Verbesserungsprozessen: schlägt Verbesserungen vor, aber zu spät und in unangemessenem Tonfall → konstruktiver vorschlagen und zu Zeitpunkt, an dem Änderungen noch möglich sind
- Erklärungen zu einzelnen Dimensionen und Verhaltensankern:
 - Formulierung: bezieht sich auf Inhalt der Sprache
 - Verständlichkeit: bezieht sich auf Form der Sprache

Beobachtungsleitfaden

- Kopfzeile ausfüllen bitte nicht vergessen. Deinen Beobachtungscode erstellst du bitte nach folgendem Schema:

▪ Vorname der Mutter:	Martina	→ MA
▪ Geburtsort:	Bamberg	→ BA
▪ Geburtstag:	25.03.1994	→ 25
Das ergibt folgenden Code: MA BA 25		

- Nicht jedes Verhalten zeigt sich bei jedem*r – Verhalten, das nicht gezeigt wird, kann auch nicht beobachtet werden.
- Notiere im Beobachtungsbogen, welches Verhalten gezeigt wird. Halte dich an die vorgegebenen Verhaltensanker.

Was hat der*die Teilnehmende im Hinblick auf...	gut gemacht:	schlecht gemacht:
Kooperation		
nonverbale Kommunikation		
Haltung		
Armhaltung (unverschränkt)	<i>dem Gegenüber zugewandter</i>	<i>verschränkte Armhaltung; dem</i>
Blickrichtung (zugewandt)	<i>Blick; Blickkontakt haltend</i>	<i>Sitznachbarn nicht den Blick</i>
Blickkontakt (aktiv suchend, haltend)		<i>zugewandt; kein Blickkontakt</i>
		<i>aktiv gesucht</i>

- Wenn Teilnehmende zum Teil gutes und zum Teil schlechtes Verhalten zeigt, vermerke das Verhalten bei gut und schlecht und konkretisiere anhand von Verhaltensankern:

Verständlichkeit		
Lautstärke (angenehm)	<i>meist angenehme Lautstärke</i>	<i>bei hitziger Diskussion zu laut</i>
Artikulation (klar, deutlich)	<i>und klare Artikulation</i>	<i>und undeutlich</i>

- Erklärungen zu einzelnen Dimensionen und Verhaltensankern:
 - Formulierung: bezieht sich auf Inhalt der Sprache
 - Verständlichkeit: bezieht sich auf Form der Sprache

Anhang G: Feedbackinstruktionen aus Studie 2

Feedbackleitfaden

- Bitte leite dein Feedback folgendermaßen ein:
 - „Wir kommen jetzt zu deinem Feedback. Vorab einige Informationen für dich:
 - Ich habe in den vergangenen 40 Minuten nur einen kleinen Teil deines Verhaltens kennengelernt.
 - Ich versuche dir das so gut wie möglich rückzumelden. Ich werde die Punkte sehr klar und deutlich formulieren, ich habe dafür standardisierte Vorgaben. All meine Aussagen sind **Vorschläge** – es ist deine Entscheidung, ob du das Feedback annimmst oder nicht.
 - Ich kann dir nur das rückmelden, was ich beobachtet habe – anhand **standardisierter Beobachtungsdimensionen**. Mehr kann ich dir nicht rückmelden.
 - Hier liegt auch Papier und Stift, wenn du möchtest, kannst du dir Notizen machen.
 - Solltest du noch Fragen oder **Gesprächsbedarf** haben, klären wir das am Ende gemeinsam mit der Studienleitung Anna-Katharina Stöcker.“
- Der Beobachtungsbogen leitet dich durch das Feedback. Formuliere die Sätze unbedingt mit beibehalten/ verändern. Das klingt dann so:
 - „Ich habe die Kompetenz ‚Kooperationsfähigkeit‘ beobachtet. Dabei habe ich als erste Dimension auf deine nonverbale Kommunikation geachtet. **Beibehalten** solltest du deine nicht verschränkte Armhaltung und deine Blickrichtung. Die war den anderen immer stark zugewandt. **Verändern** könntest du den Blickkontakt: Du kannst noch aktiver Blickkontakt suchen und dann auch halten.“
- Bei Nachfragen der Teilnehmenden, die nicht rein akustischer Art sind, reagiere folgendermaßen:
 - „Ich kann dir nur rückmelden, was ich konkret beobachtet habe. Auch in einem echten Assessment-Center sollten sich geschulte Beobachtende strikt an die Verhaltensanker halten. Am besten du stellst deine Frage gleich nochmal Anna-Katharina Stöcker, sie kann dir besser Auskunft geben.“
- Wenn das Feedback beendet ist, bitte die den*die Teilnehmende*n „Fragebogen 3“ auszufüllen.
- Wenn Teilnehmende emotional sehr aufgewühlt erscheinen, melde das bitte Anna-Katharina umgehend nach dem Feedback.

Feedbackleitfaden

- Bitte leite dein Feedback folgendermaßen ein:
 - „Wir kommen jetzt zu deinem Feedback. Vorab einige Informationen für dich:
 - Ich habe in den vergangenen 40 Minuten nur einen kleinen Teil deines Verhaltens kennengelernt.
 - Ich versuche dir das so gut wie möglich rückzumelden. Ich werde die Punkte sehr klar und deutlich formulieren, ich habe dafür standardisierte Vorgaben. All meine Aussagen sind **Vorschläge** – es ist deine Entscheidung, ob du das Feedback annimmst oder nicht.
 - Ich kann dir nur das rückmelden, was ich beobachtet habe – anhand **standardisierter Beobachtungsdimensionen**. Mehr kann ich dir nicht rückmelden.
 - Hier liegt auch Papier und Stift, wenn du möchtest, kannst du dir Notizen machen.
 - Solltest du noch Fragen oder **Gesprächsbedarf** haben, klären wir das am Ende gemeinsam mit der Studienleitung Anna-Katharina Stöcker.“
- Der Beobachtungsbogen leitet dich durch das Feedback. Formuliere die Sätze unbedingt mit gut/ schlecht. Das klingt dann so:
 - „Ich habe die Kompetenz ‚Kooperationsfähigkeit‘ beobachtet. Dabei habe ich als erste Dimension auf deine nonverbale Kommunikation geachtet. **Gut war** deine nicht verschränkte Armhaltung. Deine Blickrichtung **war zum Teil gut**. So warst du denen, die dir gegenüber saßen, sehr zugewandt. Deinen Sitznachbar hingegen hast du kaum angeschaut. **Das war schlecht. Schlecht war** außerdem dein Blickkontakt. Ich habe kaum beobachtet, dass du aktiv Blickkontakt suchst, sondern du versuchst eher, diesem auszuweichen.“
- Bei Nachfragen der Teilnehmenden, die nicht rein akustischer Art sind, reagiere folgendermaßen:
 - „Ich kann dir nur rückmelden, was ich konkret beobachtet habe. Auch in einem echten Assessment-Center sollten sich geschulte Beobachtende strikt an die Verhaltensanker halten. Am besten du stellst deine Frage gleich nochmal Anna-Katharina Stöcker, sie kann dir besser Auskunft geben.“
- Wenn das Feedback beendet ist, bitte den*die Teilnehmende*n „Fragebogen 3“ auszufüllen.
- Wenn Teilnehmende emotional sehr aufgewühlt erscheinen, melde das bitte Anna-Katharina umgehend nach dem Feedback.

Anhang H: Materialienliste aus Studie 2

Übung „Eierfall“

- 1 Bogen Zeitungspapier
- 10 Strohhalme
- 5 Büroklammern
- 5 Zahnstocher
- 1 Ei
- 1 Meter Schnur
- 1 Meter Klebeband
- 1 Schere
- 1 digitaler Küchentimer

Übung „Spaghetti-Turm“

- 2 Blatt kariertes Papier
- 1 Stift
- 50 Spaghetti
- 1 Marshmallow
- 1 Meter Schnur
- 1 Meter Klebeband
- 1 Schere
- 1 digitaler Küchentimer

Anhang I: Teilnehmendeninstruktionen aus Studie 2

Übung „Eierfall“

Rahmenbedingungen

- In dieser Übung prüfen die Beobachtenden deine Kooperationsfähigkeit sowie deine Ziel- und Ergebnisorientierung.
- Die Dauer der Übung beträgt 20 Minuten, vor euch steht eine Uhr zur Zeitkontrolle.

Aufgabe

- Vor dir steht eine Kiste mit verschiedenen Materialien.
- Ziel der Übung ist es, das Ei mit den Materialien aus der Kiste so einzupacken, dass es einen Sturz aus 2 Metern Höhe unbeschadet übersteht.
- Verwendet werden dürfen nur die Materialien aus der Kiste.

Übung „Spaghetti-Turm“

Rahmenbedingungen

- In dieser Übung prüfen die Beobachtenden deine Kooperationsfähigkeit sowie deine Ziel- und Ergebnisorientierung.
- Die Dauer der Übung beträgt 18 Minuten, vor euch steht eine Uhr zur Zeitkontrolle.

Aufgabe

- Vor dir steht eine Kiste mit verschiedenen Materialien.
- Ziel der Übung ist es, aus diesen Materialien in der Kiste einen möglichst hohen, freistehenden Turm zu bauen.
- Verwendet werden dürfen nur die Materialien aus der Kiste.
- Der Turm soll so stabil sein, dass er ein Marshmallow tragen kann. Dieses darf nur oben eingesteckt oder aufgelegt werden.
- Der Turm darf bis zum Ende der Messung nicht zusammenbrechen.
- Gemessen wird die Höhe des Marshmallows.
- Wenn dein Team im Rahmen dieser Studie den höchsten Turm baut und sich an die Aufgabenstellung hält, bekommt jedes Teammitglied einen Amazon-Gutschein über 25 Euro (für den zweithöchsten Turm gibt es je 15 Euro und für den dritthöchsten je 10 Euro).

Anhang J: Höchster Turm mit 106 cm aus Studie 2



Anhang L: Fragebogen 1 aus Studie 2

Fragebogen 1

1. Vpn-Code

Bitte fülle den folgenden Fragebogen aus. Die Daten werden nicht personenbezogen verarbeitet. Zur Zuordnung der Unterlagen benötigen wir jedoch einen personenbezogenen Code, den du bitte hier einträgst:

erste zwei Buchstaben des
Vornamens deiner Mutter

erste zwei Buchstaben
deines Geburtsortes

dein Geburtstag im
Format „Tag Tag“

Beispiel:

- Vorname der Mutter: Martina → MA
- Geburtsort: Bamberg → BA
- Geburtstag: 25.03.1994 → 25

Das ergibt folgenden Code: MA BA 25

2. Allgemeine Angaben

Datum: _____

Geschlecht: männlich weiblich keine Angabe

Studiengang: _____

(falls kein*e Studierende*r, bitte Beruf angeben)

Alter: _____

Bitte gib auf einer Skala von 1 bis 7 an, wie sehr du der folgenden Aussage zustimmst:

	1	2	3	4	5	6	7
	<small>überhaupt nicht</small>						<small>voll und ganz</small>
„Ich bekomme generell gerne Feedback.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte gib auf einer Skala von 1 bis 4 an, wie gut die einzelnen Aussagen für dich zutreffen:

	1	2	3	4
	<small>trifft gar nicht zu</small>			<small>trifft voll und ganz zu</small>
„Alles in allem bin ich mit mir selbst zufrieden.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Hin und wieder denke ich, dass ich gar nichts tauge.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich besitze eine Reihe guter Eigenschaften.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich kann vieles genauso gut wie die meisten anderen Menschen auch.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich fürchte, es gibt nicht viel, worauf ich stolz sein kann.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich fühle mich von Zeit zu Zeit richtig nutzlos.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich halte mich für einen wertvollen Menschen, jedenfalls bin ich nicht weniger wertvoll, als andere auch.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich wünschte, ich könnte vor mir selbst mehr Achtung haben.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Alles in allem neige ich dazu, mich für einen Versager zu halten.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
„Ich habe eine positive Einstellung zu mir selbst gefunden.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anhang M: Fragebogen 2 aus Studie 2**Fragebogen 2****1. Teilnehmenden-Code**

____ _ ____ _ ____ _
erste zwei Buchstaben des erste zwei Buchstaben dein Geburtstag im
Vornamens deiner Mutter deines Geburtsortes Format „Tag Tag“

2. Allgemeine Angaben

Bitte gib auf einer Skala von 1 bis 7 an, wie sehr du der folgenden Aussage zustimmst:

	1	2	3	4	5	6	7
	<small>überhaupt nicht</small>						<small>voll und ganz</small>
„Ich bin mit meiner Leistung zufrieden.“	0	0	0	0	0	0	0

Platz für Kommentare:

Anhang N: Fragebogen 3 aus Studie 2

Fragebogen 3

1. Vpn-Code

erste zwei Buchstaben des
Vornamens deiner Mutter

erste zwei Buchstaben
deines Geburtsortes

dein Geburtstag im
Format „Tag Tag“

2. Allgemeine Angaben

Bitte gib auf einer Skala von 1 bis 7 an, wie sehr du den folgenden Aussagen zustimmst:

	1	2	3	4	5	6	7
	überhaupt nicht						voll und ganz
„Ich habe dieses Mal gerne Feedback bekommen.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Mein Grund für diese Angabe: _____

„Dieses Feedback war hilfreich.“	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
----------------------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Mein Grund für diese Angabe: _____

Bitte umblättern →

Anhang O: Kodierregeln aus Studie 2

Regeln zum Rating

- Die Angaben der Vpn wurden vorab nach folgendem Schema notiert:

Vpn-Code	Angaben der Vpn (jede neue Zeile bzw. jeder Stichpunkt bzw. jeder neue Satz wurde markiert durch einen Strichpunkt)
----------	---
- Ggf. wurden auch Anmerkungen der Vpn, wie z.B. „Vielen Dank“ notiert.
- Wenn eine Vpn gar keine Vorsätze notiert hat, ist dies mit „-99“ gekennzeichnet.
- Vorgehen
 - **1. herausfiltern, was Vorsätze sind und jeden Vorsatz in einer anderen Textfarbe markieren:**
 - Als Vorsatz gilt alles, was folgenden Vorgaben nicht widerspricht
 - der Teilnehmende selbst [muss nicht genannt werden, schließt aber Formulierungen aus, in denen der Teilnehmende sagt, was andere/ „man“ tun sollte]
 - eine Verhaltensweise oder eine Eigenschaft
 - beibehalten oder verändern will [d.h. es geht nicht nur um Änderung!]
 - in zukünftigen Situationen
 - Hinweise
 - alles ist ein Vorsatz, was diesen Vorgaben nicht widerspricht, es muss aber nicht explizit genannt werden, sondern lässt sich an der Formulierung erkennen, z.B. „mehr Gestik“, „Arme nicht verschränken“
 - es zählt nicht als Vorsatz, was diesen Vorgaben widerspricht, z.B. „Eine weitere Erkenntnis ist, dass AC-Übungen nicht abschrecken müssen“
 - in einem Satz können kein, ein oder mehrere Vorsätze stecken – auch wenn die Person einen Vorsatz in Klammern mit mehreren Beispielen erläutert, können das mehrere Vorsätze sein
 - **2. Vorsätze zählen: auszählen, dabei gilt:**
 - jeder inhaltlich neue Aspekt wird gezählt, Dopplungen werden nicht gezählt
 - die Gliederung der Vpn ist nicht relevant (d.h. auch wenn Inhalte mit „und“/ „z.B.“ verknüpft sind, ist zu prüfen, ob es sich um neue Inhalte handelt)
 - im Zweifel, ob etwas eine Dopplung ist, wird der Duden herangezogen („freundlich“ ist ein Synonym von „nett“, demnach kann es herausgenommen werden)
 - wenn eine Vpn gar nichts notiert hat („-99“ im Textfeld), so zählt diese als „0“
 - **3. Auszählen der Substantive**
 - Trage in das Feld ko_nu die Anzahl der Abstraktheits-Substantive ein. Es zählen alle deutschen Substantive dazu, die folgende Endungen aufweisen:
 - heit, z. B. Gewohnheit
 - ie, z. B. Manie (aber nur, wenn es wie ein langes „i“ gesprochen wird; nicht, wenn es wie „je“ gesprochen wird, z. B. Kastanie)
 - ik, z. B. Rhythmik

- ion, z. B. Konzentration
- ismus, z. B. Protestantismus
- ität, z. B. Identität
- keit, z. B. Häufigkeit
- nz, z. B. Potenz
- tur, z. B. Struktur
- ung, z. B. Hoffnung

sowie alle Substantive der englischen Sprache, die folgende Endungen aufweisen:

- ness
 - ment
 - ship
 - dom
 - nce
 - ion
 - y (allerdings nicht, wenn es ein Diminutiv ist)
- Trage in das Feld su_nu die Anzahl der Substantive insgesamt ein.
 - Beispiele und Sonderfälle:
 - Namen zählen weder als Substantiv, noch werden sie hinsichtlich der Abstraktheit bewertet
 - Case Study: zwei Substantive, davon ein Abstraktheitssubstantiv
 - AC-Training: ein Substantiv
 - Ziel- und Ergebnisorientierung: zwei Substantive, zwei Abstraktheitssubstantive
 - AC = Assessment Center: zwei Substantive, ein Abstraktheitssubstantiv