



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Error Management Theorie und Dunning-Kruger-Effekt
– Auswirkungen auf die Einschätzung von Problemen“

verfasst von / submitted by

Mario Pérez

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magister der Naturwissenschaften (Mag.rer.nat.)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

UA 190 445 353

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UniStG
UF Biologie und Umweltkunde UniStG
UF Spanisch UniStG

Betreut von / Supervisor:

Mag^a. Dr. Sabine Tebbich, Privatdoz.

Mitbetreut von / Co-Supervisor:

Mag^a. Dr. Elisabeth Oberzaucher

Eidesstattliche Erklärung

Ich erkläre eidesstattlich, dass ich die Arbeit selbständig angefertigt, keine anderen als die angegebenen Hilfsmittel benutzt und alle aus ungedruckten Quellen, gedruckter Literatur oder aus dem Internet im Wortlaut oder im wesentlichen Inhalt übernommenen Formulierungen und Konzepte gemäß den Richtlinien wissenschaftlicher Arbeiten zitiert, durch Fußnoten gekennzeichnet bzw. mit genauer Quellenangabe kenntlich gemacht habe.

Wien, am 29.03.2020

Mario Pérez

Danksagung

Zum Erstellen dieser Abschlussarbeit bedurfte es der Unterstützung vieler Leute, denen ich hiermit allen meinen Dank aussprechen möchte:

Ich möchte mich zuallererst bei meiner Betreuerin, Frau Mag^a Dr. Sabine Tebbich für die Unterstützung und gute Betreuung beim Verfassen dieser Arbeit bedanken.

Ein großer Dank gebührt meiner Mentorin und Mitbetreuerin Frau Mag^a Dr. Elisabeth Oberzaucher, welche immer mit großer Hilfsbereitschaft, guten Tipps und nicht enden wollender guter Laune, aber auch mit mahnender Souveränität hinter mir stand.

Weiters möchte ich meiner Familie, vor allem aber meiner Mutter danken, ohne deren Rückhalt und Motivation diese Arbeit wohl (noch) längere Zeit in Anspruch genommen hätte.

Meinen Freunden, allen voran dem Bio Stammtisch und meinen beiden Mitbewohnern, weiters den Mitgliedern der Familie Meier, welche mich in allen Fragen auf das Herzlichste unterstützt haben. Ohne euch wäre der Weg hierher um einiges grauer gewesen und ich bin froh, dass ihr mich durch das Studium und darüber hinaus begleitet habt.

Zu guter Letzt gilt mein Dank auch den 302 Personen, die freiwillig und gratis, aber nicht umsonst an meiner Studie teilgenommen haben.

Abstract

Menschen bedienen sich bei der Einschätzung von Problemen mehrerer Hilfsmittel, welche einerseits Resultate der Evolution sind, andererseits aber auch in der persönlichen Entwicklung des Menschen entstehen können. Ein Beispiel hierfür ist die Error Management Theory (EMT), welche von Haselton und Buss im Jahre 2000 entwickelt worden ist und auf der Theorie des Asymmetrischen Minimalinvestments nach Trivers (1972) basiert. Sie besagt, dass Frauen als das Geschlecht mit einem größeren biologischen Minimalinvestment zu einem Unterschätzen des männlichen Investmentbestrebens tendieren sollten, während Männer das Reproduktionsinteresse von Frauen tendenziell überschätzen sollten. Weiters können verschiedene Effekte Einflüsse auf die Selbsteinschätzung von Personen haben, unter anderem der Dunning-Kruger-Effekt (1999). Dieser sagt aus, dass Menschen mit weniger Erfahrung die Tendenz zeigen, sich zu überschätzen, während Menschen mit mehr Erfahrung sich eher unterschätzen.

Anhand eines Feldversuches mit zwei Geduldsspielen wird untersucht, ob die EMT in einem nicht-reproduktiven Kontext wirksam und ob ein Dunning-Kruger-Effekt bei den Personen beobachtbar ist. Dazu sind 302 Personen mithilfe eines Fragebogens und der beiden Geduldsspiele befragt worden. Es kann bei den männlichen Teilnehmern eine Überschätzung im Sinne der EMT dokumentiert werden, wobei die Erfahrung der Teilnehmer*innen keine signifikante Auswirkung auf die Selbsteinschätzung zeigt, jedoch aber das Befinden.

Humans use several tools for the assessment of problems, which on the one hand are results of evolution, but on the other hand can also arise in the personal development of humans. An example is the Error Management Theory (EMT), which was established by Haselton and Buss in 2000 and is based on the theory of Asymmetric Minimal Investment according to Trivers (1972). It states that women, as the sex with a larger biological minimum investment, tend to underestimate the male investment effort, while men tend to overestimate the reproductive interest of women. Furthermore, various effects can have an influence on the self-assessment of individuals, including the Dunning-Kruger effect (1999). This states that people with less experience tend to overestimate themselves, while people with more experience tend to underestimate themselves.

A field trial with two puzzle games was used to investigate whether EMT is effective in a non-reproductive context and whether a Dunning Kruger effect can be observed in the individuals. For this purpose, 302 persons were interviewed by means of a questionnaire and the two puzzle games. An overestimation in the sense of EMT could be documented among the male participants, while the experience of the participants did not show a significant effect on self-assessment, the well-being however, did.

Inhaltsverzeichnis

1. Einleitung	7
1.1 EMT – Die Error Management Theory	7
1.1.1 EMT im nicht-reproduktiven Kontext	8
1.1.2 EMT im reproduktiven Kontext	10
1.2 Minimal Investment nach Trivers und das Bateman-Prinzip	11
1.3 Resümee EMT und Minimal Investment	13
1.4 Die „Illusory superiority“	14
1.4.1 Der Primus-inter-pares-Effekt	14
1.4.2 Der Better-than-average-Effekt	15
1.4.3 Der Worse-than-average-Effekt	18
1.5 Der Dunning-Kruger-Effekt	19
1.5.1 Die Studie zum Dunning-Kruger-Effekt	19
1.5.2 Erklärungen für den Dunning-Kruger-Effekt	21
1.6 Hypothesen	23
2. Methoden	25
2.1 Anzahl an Versuchsteilnehmer*innen	25
2.2 Auswahl Standorte	25
2.3 Vorgehen bei der Datenerhebung	25
2.4 Fragebogen	26
2.5 Die Geduldspiele	28
2.6 Auswertung	28
3. Ergebnisse	29
3.1 Selbsteinschätzung nach der Error Management Theory	29
3.1.1 Puzzlespiel „Ring mit Stern“	29

3.1.2 Legespiel „Die verflixten Spechte“	31
3.2 Selbsteinschätzung nach dem Dunning-Kruger-Effekt	32
3.2.1 Puzzlespiel „Ring mit Stern“	32
3.2.2 Legespiel „Die verflixten Spechte“	34
3.3 Weitere interessante Ergebnisse	37
4. Diskussion	39
4.1 Interpretation der Selbsteinschätzung nach der EMT	39
4.2 Interpretation der Selbsteinschätzung nach Dunning-Kruger	40
4.3 Ausblick auf künftige Studien	42
5. Literatur	44
Internetquellen	47
6. Anhang	48
Fragebogen	48

1. Einleitung

In dieser Abschlussarbeit für das Diplom-Lehramt Biologie und Spanisch werden die Einflüsse der Error Management Theorie von Haselton und Buss sowie des Dunning-Kruger-Effektes auf die Selbsteinschätzung von Menschen beim Lösen von Problemen untersucht.

Die theoretischen Grundlagen zu den genannten Theorien sollen hier nun angeführt und für die Leser*innen erläutert werden, um ein Verständnis für die Thematik, die angewandten Methoden und den Ablauf des Experimentes zu generieren. Dieses wird im Anschluss erklärt, so wie auch die Ergebnisse der praktischen Untersuchungen präsentiert werden. Abschließend werden diese Ergebnisse diskutiert und ein kurzer Ausblick für weitere Forschungen gegeben.

1.1 EMT – Die Error Management Theory

Im Jahr 2000 beschreiben Martie G. Haselton und David M. Buss aufgrund ihrer Untersuchungen erstmals näher eine Theorie, welche sich mit dem Treffen von Entscheidungen beim Menschen auseinandersetzt: Die Error Management Theory besagt laut den beiden Autor*innen, dass kognitive Fehler aus adaptiven Vorurteilen resultieren, die in der Gegenwart existieren, weil sie in der Vergangenheit zu Überlebens- und Reproduktionsvorteilen für den Menschen geführt haben (Haselton, Buss, 2000). Für die Autor*innen hat sich also das Treffen von Entscheidungen im Laufe der natürlichen und sexuellen Evolution entwickelt und verändert, da sich Menschen bei diesen Vorgängen auf ihr Vorwissen, sei es erlernt oder von Vorfahren mitgegeben, aber auch auf Vorurteile über gewisse Fakten und Umstände verlassen mussten und heute immer noch müssen, um in bestimmten Situationen die (für sie) richtigen Handlungen zu setzen.

Dabei können Personen nun, so Haselton und Buss, zwei Arten von Fehlern begehen, so die Entscheidung über einen Aspekt überhaupt als „Fehlentscheidung“ bezeichnet werden kann:

- Einerseits ein sogenanntes „false-positive“ oder „Typ-1-Fehler“; hierbei handelt es sich um eine falsche Einschätzung einer Situation, die so nicht stattgefunden hat,

beispielsweise bei einem Feuermelder, der eine Warnung ausgibt, obwohl kein Feuer ausgebrochen ist.

- Andererseits ein „false-negative“ oder „Typ-2-Fehler“, wobei hier eine Situation nicht richtig erkannt worden ist und somit auch keine adäquate Handlung gesetzt werden hat können. Ein Beispiel hierfür ist ein Feuermelder, welcher ein Feuer nicht detektiert hat und somit keine Warnung ausgegeben hat (Haselton, Buss, 2000).

Wie bereits an den gegebenen Beispielen zu erkennen ist, werden diese Prinzipien des Error Managements nicht nur zur Erklärung des menschlichen Verhaltens herangezogen, sondern auch zum Beispiel in der Technik, in der Medizin und in anderen für den Menschen relevanten Bereichen.

1.1.1 EMT im nicht-reproduktiven Kontext

Anhand eines weiteren Musterfalls werden nun vier Möglichkeiten veranschaulicht, die beim Treffen einer Entscheidung eintreten können: Trifft ein Mensch in der Natur auf einen schlangenförmigen Gegenstand, so muss er ein Urteil darüber fällen, ob es sich bei besagtem Objekt um eine Schlange oder einen Ast handelt. Glaubt er, es handle sich um eine Schlange, und es liegt tatsächlich eine Schlange vor ihm, so tritt ein *true-positive*-Fall ein. Liegt jedoch stattdessen ein Ast vor ihr, so hat die Person einen Typ-1-Fehler (*false-positive*) begangen. Umgekehrt ist es der Fall, wenn die Person annimmt, der Gegenstand vor ihr sei ein Ast, und dies der Tatsache entspricht: Dann wird dies als ein *true-negative*-Fall bezeichnet. Der Typ-2-Fehler (*false-negative*) wird dann begangen, wenn die Person annimmt, einen Ast vor sich zu haben, es sich aber hingegen um eine Schlange handelt. Ob das Ergebnis nun positiv oder negativ für die betroffene Person ist, hängt immer von der Art der Situation und der „Bedrohung“ ab, in diesem Fall, ob die Schlange richtig oder falsch erkannt worden ist (Galperin, Haselton, 2013).

Das nachfolgende Modell soll die vier Entscheidungsmöglichkeiten in einer einfachen Tabelle aufzeigen:

		Annahme	
		Schlange	Ast
Tatsächlicher Zustand	Schlange	True-positive	False-negative
	Ast	False-positive	True-negative

Tabelle 1: Möglichkeiten beim Treffen einer Entscheidung im Rahmen der EMT (verändert nach Galperin, Haselton, 2013)

Wichtig ist dabei zu erwähnen, dass ein getroffener Entschluss, sollte er sich als Fehler herausstellen, nicht unbedingt eine falsche Entscheidung sein muss: Je nach Situation belaufen sich die Kosten eines Fehlers, also die daraus resultierenden Konsequenzen, nämlich auf unterschiedliche Wertigkeiten, so David M. Green und John A. Swets (1966). Auf diese beiden Autoren beziehen sich auch Haselton und Buss bei der Erhebung ihrer Error Management Theory (Haselton, Buss, 2000, 2009). Nach deren Signalentdeckungstheorie (Green, Swets, 1966) ist meist einer der beiden zu machenden Fehler mit einem größeren Kostenaufwand verbunden. Des Weiteren kann jeweils nur einer der beiden Fehler minimiert werden. Verringert man die Wahrscheinlichkeit eines Fehlers, so steigt korrelativ dazu die Wahrscheinlichkeit des anderen Fehlers. Aufgrund dieser bivalenten Verteilung und Abhängigkeit der Fehler-Arten gibt es zwei Arten von Systemen: Jene, bei denen ein Typ-1-Fehler kostspieliger ist, wie beispielsweise bei der Testung von wissenschaftlichen Hypothesen (Haselton, Buss, 2000), und jene, bei denen ein Typ-2-Fehler gravierendere Konsequenzen hervorruft; Beispiele hierfür sind unter anderem die beiden oben genannten, der Feuermelder und das Schlangendilemma. Bei Letzterem wäre es für die Person weitaus schlechter, würde sie eine Schlange für einen Ast halten, da damit die Gefahr, gebissen zu werden, viel höher wäre. Die Kosten beliefen sich hier auf den wahrscheinlichen Tod und somit auch auf den Verlust des Reproduktionserfolges, wohingegen sich die Kosten im Falle eines Typ-1-Fehlers (einen Ast für eine Schlange halten) auf eine erhöhte Vorsicht beliefen. Genauso verhält es sich bei dem Feuermelder-Beispiel. Diese Geräte werden in der heutigen Produktion bereits übersensibilisiert hergestellt, um eher in Richtung eines Typ-1-Fehlers zu melden. Somit ist es möglich, die

Kosten dieses Fehlers im Falle einer Falschmeldung niedrig zu halten. Diese beziehen sich auf das „überflüssige“ Alarmieren der betroffenen Personen im Gegensatz zu einem nicht detektierten Feuer und den damit verbundenen, potentiell höheren Kosten des möglichen Versehrens und Verlustes von menschlichem Leben und materiellem Hab und Gut (Nesse, 2001).

Tendenziell geht die Forschung davon aus, dass Systeme sich immer für die kostengünstigere Variante entscheiden, wenn die Fehlerkostenverteilung asymmetrisch ist (Haselton, Buss, 2000). Legt man diese Überlegungen nun auf das menschliche Verhalten um, so lässt sich im Laufe der Evolution beobachten, dass auch der Nutzen asymmetrisch verteilt ist. Die Error Management Theory besagt, der Mensch habe sich dazu entwickelt, jene Voreingenommenheitsregeln anzuwenden, welche ein Ergebnis von höherem Nutzen oder aber von niedrigeren Fehlerkosten lieferten, selbst wenn diese Entscheidungsregeln häufiger Fehler produzierten (Haselton, Buss, 2000). Die Voraussetzung hierbei ist, dass die alternativen Entscheidungsregeln noch kostspieliger sind.

1.1.2 EMT im reproduktiven Kontext

Ein veranschaulichendes Szenario im reproduktiven Kontext ist folgendes: Bei der Entscheidung zwischen einer langanhaltenden Partnerschaft oder einem kurzzeitigen Dating müssen Menschen die jeweiligen Vor- und Nachteile abwägen. Da zum Beispiel die Kosten der Fortpflanzung für Frauen höher als für Männer sind, wird konstatiert, dass Frauen sowohl bei der Wahl ihrer Partner für kurzfristige als auch für langfristige Beziehungen generell wählerischer seien. Es wird überdies angenommen, dass Frauen im Durchschnitt kurzfristige Beziehungen weitgehend vermeiden, da die potenziellen Kosten gegenüber den potenziellen Vorteilen überwiegen. Die potenziellen Kosten sind in diesem Fall das generelle Risiko, schwanger zu werden, wobei nicht sicher davon auszugehen ist, dass der Partner sich zur Aufzucht des Kindes verpflichtet fühlt, oder das erhöhte Risiko einer sexuell übertragbaren Krankheit. Bei Männern kann hinsichtlich dieser Thematik grundsätzlich das Gegenteil angenommen werden, solange die potenziellen Vorteile (Chance auf Reproduktion und wenig Energieaufwand) die potenziellen Nachteile (eben jener Energie-/ Ressourcenaufwand und mögliches Risiko einer sexuell übertragbaren Krankheit) einer kurzzeitigen Beziehung übersteigen (Geary, Vigil, Byrd-Craven, 2004).

Diese Erkenntnisse über die Kosten/Nutzen-Abwägung in Populationen postuliert Robert Trivers bereits 1972 und benennt seine Forschungsergebnisse mit dem Titel des Minimalinvestments in der Elternschaft. Sie sollen im Folgenden nun kurz beschrieben werden.

1.2 Minimal Investment nach Trivers und das Bateman-Prinzip

Ausgehend von Beobachtungen des Genetikers Angus John Bateman, welcher die Reproduktionserfolge von Männchen und Weibchen bei Fruchtfliegen der Art *Drosophila melanogaster* studiert hat (1948), entwickelt Trivers die Theorie des elterlichen Investments. Er definiert dieses wie folgt: "I first define parental investment as any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and hence reproductive success) at the cost of the parent's ability to invest in other offspring." (Trivers, 1972, S. 55).

Jeder energetische Aufwand also, den ein Elternteil aufbringt, um das Leben und Überleben des eigenen Nachwuchses zu sichern, gilt bei Trivers als Investment. Dazu zählen sowohl körperliche Vorgänge wie die intrauterine Ernährung eines Fötus oder die Laktation bei Frauen als auch die Aufzucht und Fürsorge der Nachkommen, welche zeitintensiv und somit auch ressourcenverbrauchend sein können (Trivers, 1972). Zeit, welche nun ein Elternteil nicht in die Zeugung weiterer Nachkommen investieren kann. Dieser Faktor wirkt sich freilich bei Frauen ob der vorhandenen und unterschiedlich langen Tragzeit des Nachwuchses naturgemäß stärker aus als bei Männern.

Bateman stellt nach den Ergebnissen seiner Experimente fest, dass auf die verschiedenen Geschlechter ein unterschiedlicher Selektionsdruck erfolge. Bei den Männchen handle es sich hierbei um eine intrageschlechtliche Selektion, da bei ihnen eine höhere Fruchtbarkeit abhängig ist von einer erhöhten Häufigkeit, also Frequenz an Inseminationen (Bateman, 1948). Die Männchen unterliegen also, im Gegensatz zu den Weibchen dieser Art, einem Selektionsdruck gegenüber ihren gleichgeschlechtlichen Mitbewerbern, gegen welche sie sich durchsetzen müssen, um ihre genetische Erbinformation weitergeben zu können. Erklärbar ist dieses Phänomen laut Bateman mit dem ungleich großen Wachstum und der ungleichen Verteilung der Keimzellen bei den Geschlechtern, der Anisogamie (Bateman, 1948). Die männlichen Gameten sind hierbei meist kleiner und begeißelt, also zu einer

Fortbewegung befähigt, die weiblichen Gameten sind meist um ein Vielfaches größer und unbegeißelt und deshalb nicht aktiv zu einer Bewegung fähig. Da die Anisogamie schon sehr bald in der Evolution aufgetreten ist, argumentiert er weiters, dass dies auf fast alle Lebewesen umsetzbar sei, ausgenommen hiervon seien einige primitive Lebensformen und Organismen, welche monogam leben und einer Geschlechterverteilung von 1:1 und somit keiner intra-geschlechtlichen Konkurrenz unterliegen (Bateman, 1948).

Trivers nimmt in seiner Theorie diese Erklärung auf: Beim Menschen muss der Mann viel weniger Energie in die Produktion seiner Gameten investieren (die Samenzellen sind klein, meist kurzlebig und unbegrenzt reproduzierbar), wohingegen bei der Frau das Energieinvestment in die weiblichen Gameten um einiges höher ist, da die Zellen einerseits größer sind und andererseits auch länger überlebensfähig sein müssen (Trivers, 1972). Dazu kommt, dass der Vorrat an weiblichen Eizellen schon von der Geburt der Frau an begrenzt ist, sie also nur eine eingeschränkte Anzahl von ca. 400 reproduktionsfähigen Gameten in ihrem Lebenszyklus besitzt. Daraus folgt für Trivers nun, dass die Frau das Geschlecht mit dem höheren biologischen Minimalinvestment in ihren Reproduktionserfolg beziehungsweise in ihren Nachwuchs ist (Trivers, 1972). Dies bewahrheitet sich dazu auch noch nach der erfolgreichen Fortpflanzung, berücksichtigt man das weitere energetische Investment einer Frau in ihre Nachkommen während der neunmonatigen Schwangerschaftsperiode, in Form der anhaltenden Produktion von Muttermilch nach der Geburt und bei der anschließenden Aufzucht. Für den Mann hingegen ist das minimale Investment, welches er in die Weitergabe seiner Gene einbringen muss, theoretisch nach der erfolgreichen Kopulation erfolgt und er kann, muss jedoch nicht zwingend, weitere Energien mobilisieren, um das Überleben seiner Nachkommen zu sichern (Trivers, 1972). Die Selektion wirkt also auf die beiden Geschlechter unterschiedlich:

- Bei Männern ist demnach nicht die Menge an produzierten Gameten der limitierende Faktor für ihren Reproduktionserfolg, sondern ihr Erfolg, mit diesen reifen Geschlechtszellen weibliche Eizellen zu befruchten, ergo ihre effektive, positive Bilanz beim weiblichen Geschlecht. Dies führt naturgemäß zu einer Konkurrenz der Männer um paarungsbereite Weibchen (die „limitierende Ressource“).

- Bei Frauen verhält sich der Reproduktionserfolg umgekehrt proportional: Sie sind nicht limitiert durch die erfolgreiche Befruchtung ihrer Eizellen, sondern durch die Fähigkeit, nur eine begrenzte Anzahl geschlechtsreifer Gameten zu produzieren.

Infolgedessen scheinen Frauen bei der Auswahl ihrer Sexualpartner wählerischer zu sein und auf Merkmale bei Männern achten, welche von einer erhöhten genetischen Qualität zeugen (Bateman, 1948; Trivers, 1972).

1.3 Resümee EMT und Minimal Investment

Verknüpft man nun die beiden Theorien von Haselton und Buss (Error Management Theory, 2000) und Trivers (Minimal Investment, 1972), kann Folgendes zusammenfassend gesagt werden, natürlich nun im Hinblick auf den reproduktiven Kontext: da auf Männern ein innergeschlechtlicher Konkurrenzdruck um die limitierende Ressource „Frau“ herrscht, sollten sie dazu tendieren, sich optimistischer in Bezug auf ihren Erfolg bei Frauen einzuschätzen, also einen Typ-I-Fehler zu begehen, da die Kosten einer nicht wahrgenommenen Reproduktionschance höher sind als die Aufwandskosten, in welche eine Abfuhr von Seiten des weiblichen Parts resultieren würden. Bei Frauen verhält sich dies den Theorien nach umgekehrt. Diese sollten aufgrund der natürlichen Limitierung ihrer Geschlechtszellen eher zu einem Typ-II-Fehler tendieren, also sich selbst pessimistischer einschätzen und Sexualpartner gründlicher auswählen, da die Kosten für sie höher wären, würden sie einen Partner wählen, welcher im Falle einer erfolgten Reproduktion nicht für sie sorgen würde beziehungsweise könnte, als wenn sie eine Reproduktionschance nicht wahrnehmen.

Zusätzlich muss natürlich berücksichtigt werden, dass die soziale Komponente beim Menschen hinsichtlich dieser biologischen Theorie nicht vernachlässigt werden kann. Die gesellschaftlichen Umstände unseres Sozialsystems ermöglichen es dem Menschen seit dem späten 20. Jahrhundert, die Lasten der Investments aufzuteilen, so etwa bei der Kinderpflege durch Dritte oder der Spaltung der erzieherischen Maßnahmen auf Eltern, Schulsystem und soziale Interaktionen. Außerdem spielen die genetischen Faktoren und Vorbestimmungen nicht mehr eine so prägnante Rolle für das Überleben der Kinder als noch zu Zeiten eines weniger stark ausgeprägten Gesundheitssystems. Trotz der oben angesprochenen Faktoren, die im Laufe der Zeit einer Veränderung unterworfen worden sind, sind die Theorien von Bateman und Trivers von nicht unbeachtlicher Relevanz für die Untersuchung, die dieser Diplomarbeit zugrunde liegt.

1.4 Die „Illusory superiority“

Menschen tendieren generell dazu, sich selbst eine bessere Bewertung zu geben, wenn es darum geht, Einschätzungen zu eigenen Kompetenzen und Fähigkeiten abzugeben. Untersuchungen auf diesen Gebieten der sozialen Kognition und der Sozialpsychologie gibt es viele (Codol, 1975; Chi, 1978, 1982; Alicke, 1985; Dunning, Meyerowitz, Holzberg, 1989; Van Yperen, Buunk, 1991), zusammengefasst worden sind die Entdeckungen der Autor*innen unter dem Namen der „*illusory superiority*“ (zu Deutsch etwa: „Scheinbare Überlegenheit“). Diesen Begriff verwenden zuallererst Nico W. Van Yperen und Bram P. Buunk im Jahre 1991, welche unter anderem Forschungen zu den Geschlechterrollen in Beziehungen und zur Zufriedenheit in Ehebündnissen angestellt haben (Van Yperen, Buunk, 1990). Im Zuge dieser Forschungen haben sie feststellen können, dass viele Paare dazu tendieren, die Zufriedenheit ihrer Beziehung anhand ihres jeweiligen Investments in die Beziehung eher schlechter einzuschätzen, jedoch besser im Vergleich zu Beziehungen anderer Paare (Van Yperen, Buunk, 1991). Ein Effekt, den Buunk später im Jahre 1997 salopp damit übertitelt, dass „die meisten Beziehungen gut sind, unsere aber die Beste ist“ (Buunk, van der Eijnden, 1997, S. 219). Sie können also beweisen, dass Menschen scheinbar allzu oft davon ausgehen, besser als andere oder zumindest der Durchschnitt der Bevölkerung zu sein, vor allem, wenn sie in den jeweiligen Bereichen weniger Erfahrung haben. Interessant ist dabei, dass man diese Beobachtung gewissermaßen auf viele Bereiche anwenden kann.

1.4.1 Der Primus-inter-pares-Effekt

Bereits im Jahre 1975 erhebt Jean-Paul Codol anhand von insgesamt zwanzig Experimenten Daten, welche darauf rückschließen lassen, dass einzelne Personen in einer Gruppe behaupten, sich stärker an die in der Gruppe geltenden Normen zu halten als alle anderen Mitglieder dieser Gruppe. Dieses Phänomen wird als „*superior conformity of the self phenomenon*“ (zu Deutsch etwa: „Phänomen der Überlegenheitskonformität des Selbst“) oder *Primus-inter-pares*-Effekt (PIP-Effekt), also den „Erster-unter-Gleichen-Effekt“, bezeichnet (Codol, 1975). Dieser Effekt stellt eine grobe Basis für alle späteren Untersuchungen zu dieser Eigenschaft dar und wird oft in der Literatur angegeben.

Unter anderem zum Beispiel von der Kognitions- und Erziehungswissenschaftlerin Michelene T. H. Chi. In ihrer Forschung zur Lernentwicklung von Kindern beobachtet sie, dass Kinder als Schachneulinge längere Zeit und mehr Züge benötigen, um die Positionen verschiedener Spielfiguren auf einem Schachbrett zu rekonstruieren als Spieler*innen mit mehr Erfahrung, wobei jedoch die Schachneulinge vorher die Anzahl an Spielzügen niedriger eingeschätzt haben (Chi, 1978).

1.4.2 Der Better-than-average-Effekt

Bezogen auf Codol wird dieses Eigenschaftsmerkmal in der Studie von Mark D. Alicke im Jahre 1985 und in jener von Alicke und Olesya Govorun aus dem Jahre 2005 untersucht. Alicke lässt College-Student*innen sich selbst in Bezug auf verschiedene Merkmale einschätzen und bringt in Erfahrung, dass diese Einschätzung im Vergleich zum Durchschnitt der befragten Student*innen immer positiver ausgefallen ist, vor allem dann, wenn die Merkmale in ihrer positiven Konnotation steigen (Alicke, 1985). Dies geschieht laut dem Autor vor allem auch aus dem Grund, um das eigene Selbstwertgefühl und das Selbstbewusstsein zu steigern (Alicke, 1985).

In einer Metastudie fassen Alicke und Govorun im Jahre 2005 einige der Ergebnisse der vorhergehenden Forschungen zusammen und erklären somit den sogenannten „*Above-average-Effekt*“ oder auch „*Better-than-average-Effekt*“ (Alicke, Govorun, 2005). Auch dieser charakterisiert das in den verschiedenen Disziplinen vorliegende Phänomen, dass sich Personen grundsätzlich besser als der Durchschnitt einschätzen. Das Ziel ihrer Studie ist es, verschiedene Erklärungsmöglichkeiten des *Better-than-average-Effektes* zu sammeln. Laut den zwei Autor*innen gibt es mehrere dieser Erklärungen, unter anderem folgende:

Linda S. Perloff und Barbara K. Fetzner untersuchen 1986, wie sich Personen, die keine Opfer negativer Lebensereignisse sind, mit anderen Personen vergleichen, und zwar mit dem Durchschnitt jener Werte. Dabei finden sie heraus, dass diese sich selektiv mit Menschen vergleichen, welche gerade Opfer negativer Ereignisse gewesen sind, um, wie die Autorinnen behaupten, eben im Vergleich besser abzuschneiden (Perloff, Fetzner, 1986).

Umgekehrt kann Neil D. Weinstein im Jahr 1980 beweisen, dass Personen selektiv Merkmale heranziehen, in denen sie besser als ihre Mitmenschen sind, um denselben Effekt zu erzielen, nämlich im Vergleich besser dazustehen (Weinstein, 1980). Er lässt in seiner ersten Studie 300 College-Student*innen ihre Chancen, dass bestimmte Events in ihrer Zukunft eintreten könnten in 42 verschiedenen Ereignissen im Vergleich zu ihren Kolleg*innen einschätzen. Zu diesen Events gehören zum Beispiel das Reisen nach Europa, der Besitz eines eigenen Hauses oder die Möglichkeit, Lungenkrebs zu bekommen. In einer zweiten Studie lässt er die Student*innen Faktoren aufzählen, welche ihre Chancen erhöhen würden, positive zukünftigen Events herbeizuführen oder negative Events zu vermeiden (Weinstein, 1980). Die Ergebnisse betitelt er als „*unrealistic optimism*“, also zu Deutsch „unrealistischer Optimismus“, beziehungsweise als „*optimistic bias*“ („optimistische Voreingenommenheit“). Dieser Optimismus entstehe laut Weinstein dadurch, dass die Versuchspersonen so sehr auf jene bestimmten Faktoren achteten, welche ihre eigenen Chancen auf ein positives Outcome verbesserten, sodass sie außer Acht ließen, dass andere dieselben Faktoren ebenfalls besitzen könnten (Weinstein, 1980). Diese optimistische Voreingenommenheit existiert auch, wenn Menschen die Wahrscheinlichkeit von zukünftigen positiven Ereignisse schätzen sollen. Das Eintreffen von für den Menschen positive Ereignisse wird statistisch gesehen eher überschätzt, wohingegen die Wahrscheinlichkeit für eintretende negative Ereignisse unterschätzt wird (Sharot, 2011).

Als Resümee der verschiedenen Erklärungsmöglichkeiten des *Above-average*-Effektes soll nun noch einmal auf die Metastudie von Alicke und Govorun vom Jahre 2005 Bezug genommen werden. Sie gliedern die verschiedenen Erklärungsmöglichkeiten in insgesamt fünf Unterpunkte (Alicke, Govorun, 2005):

- „Die Selektive Rekrutierung“: Mit diesem Begriff fassen die Autor*innen unter anderem die Erkenntnisse von Perloff und Fetzer von 1986 sowie von Weinstein 1980 zusammen, woraufhin Personen, wenn sie sich selbst mit den durchschnittlichen Ergebnissen einer Gruppe vergleichen sollen, dazu tendieren, selektiv Merkmale auszusuchen, also diese zu „rekrutieren“, um vergleichsweise besser dazustehen. Dies kann auf zwei Arten geschehen: Die Auswahl der Kriterien erfolgt entweder nach eigenen, positiven Merkmalen oder aber nach negativen Merkmalen von anderen.

- „Der Egozentrismus“: Bezugnehmend auf vor allem die Arbeit von Kruger im Jahre 1999, welcher den Begriff des egozentristischen Urteiles in seiner Forschung prägt (Kruger, 1999), legen Alicke und Govorun diese Erklärung folgendermaßen dar: Bei einer Selbsteinschätzung von Menschen gehen diese bevorzugt von den eigenen Erfahrungen, Überzeugungen und Werten aus und beurteilen diese besser als die ihrer Vergleichsgruppe respektive den Durchschnitt dieser. Der ausschlaggebende Unterschied zur selektiven Rekrutierung liegt nun laut den Autor*innen im Faktum, dass bei der egozentristischen Erklärungsweise der Auswahl der Verhaltensmerkmale keine Eigennützigkeit zugrunde liegt (Alicke, Govorun, 2005). Die Items der Vergleichsgruppe werden in jedem Falle anerkannt, allerdings werden die eigenen Charakteristika als positiver wahrgenommen und eingestuft. Interessant hierbei ist, dass dem egozentristischen Erklärungsversuch die *Theory of mind* zu einer notwendigen Grundbedingung gemacht werden kann. Diese Theorie im vollen Umfang auszuführen, würde den Rahmen der Diplomarbeit sprengen, es sei jedoch so viel dazu gesagt: Diese Theorie beschreibt die Fähigkeit eines Menschen, sich selbst gewisse psychische Konzepte wie Überzeugungen, Wissensstand, Emotionen und Absichten zuzuschreiben und aber auch zu erkennen, dass andere ebenso diese Prinzipien haben und diese sich von den eigenen unterscheiden können (Förstl, 2012).
- „Der Fokalismus“: Wie ausgeprägt der *Above-average*-Effekt bei diesem Erklärungsansatz ist, hängt laut David A. Schkade und Daniel Kahnemann davon ab, von welchem Standpunkt aus die Einschätzung erfolgt (Schkade, Kahnemann, 1998). Sie haben untersucht, wie glücklich sich Einwohner*innen Kaliforniens im Vergleich zu anderen Bewohner*innen der USA, namentlich des Mittleren Westens, einschätzen. Wenn der Fokus der Bewertung auf die eigene Person im Vergleich zum Durchschnitt liegt, tritt der *Above-average*-Effekt deutlich hervor. Umgekehrt oder zumindest vermindert sollte der Effekt nun auftreten, wenn die Vergleichsgruppe in ihrer Gesamtheit als Referenz zum Selbst in den Fokus gelegt wird (Alicke, Govorun, 2005). Der Fokalismus wird vor allem auch durch die *optimistic bias* gestützt und ist damit einhergehend: In verschiedenen Studien (Otten, van der Pligt, 1966; Eiser, Pahl, Prins, 2001) wird gezeigt, dass die optimistische Voreingenommenheit höher ist, wenn der Durchschnitt die Referenz stellt und die Proband*innen sich selbst einschätzen.

- „Der Vergleich des Selbst mit Gruppierungen“: Die Autoren Yechiel Klar und Eilath E. Giladi haben unter anderem erforscht, welche Auswirkungen die oben beschriebenen Effekte haben, wenn eine Person eine ihr unbekannte, einzelstehende Person im Vergleich zu einer Gruppierung bewertet (Klar, Giladi, 1997). Sie legen dar, dass diese Einzelstehenden besser bewertet werden als der Durchschnitt der Gruppe. Auffallend ist, dass dieses Merkmal, wie auch schon bei vorher angeführten Beispielen, einen Effekt sowohl ins Positive als auch ins Negative zeigt: Die einzelne Person aus einer als positiv angesehenen Gruppierung wird über dem Durchschnitt bewertet, jedoch unter dem Durchschnitt, wenn die Gruppierung als solche als negativ angesehen wird. Weiters stellen sie fest, dass diese Effekte auch auftreten, wenn materielle, nichtmenschliche Objekte wie Seifendüfte oder Musikauswahlmöglichkeiten bewertet werden. Für Alicke und Govorun liegt dieses Erkenntnis dem *Above-average*-Effekt zugrunde (Alicke, Govorun, 2005).

1.4.3 Der *Worse-than-average*-Effekt

Erweiternd zum *Better-than-average*-Effekt liefert Justin Kruger im Jahre 1999 in seiner Arbeit zur egozentristischen Natur des vergleichenden Urteils den Beweis, dass dieses Phänomen auch in gegensätzlicher Auswirkung vorliegt (Kruger, 1999). Er geht von der Annahme aus, dass Menschen sich entsprechend dem *Above-average*-Effekt überschätzen, wenn sie von sich ausgehend (also egozentristisch) ihren Stand im Vergleich zum Durchschnitt bewerten. Dies geschieht vor allem dann, wenn es um Aufgaben oder Merkmale geht, die sie als „einfach“ definieren. Also, so folgert Kruger, müssten sich Personen nun unterschätzen, wenn sie den Wert und die Erreichbarkeit von für sie „schweren“ Aufgaben oder Merkmalen gegenüber dem Durchschnitt egozentristisch ansehen (Kruger, 1999). Er bezeichnet diesen Effekt als „*Worse-than-average*-Effekt“ oder „*Below-average*-Effekt“. In zwei Studien untersucht er diese Theorie, indem er unter anderem Student*innen einige Fähigkeiten nach ihrer Attraktivität beziehungsweise ihrem Wert des Erstrebens anordnen sowie ihren eigenen und den Level ihrer Kolleg*innen in den jeweiligen Fähigkeiten angeben lässt. Tatsächlich unterschätzen sich die

Studienteilnehmer*innen in eben jenen Fähigkeiten, die sie als weniger erstrebenswert oder auch für sie schwerer erreichbar ansehen (Kruger, 1999).

1.5 Der Dunning-Kruger-Effekt

Der Dunning-Kruger-Effekt wird von David Dunning und Justin Kruger im Jahre 1999 erstmalig beschrieben. Sie beobachteten, in Einverständnis mit der vorangegangenen Forschung rund um die „*illusory superiority*“, dass Menschen im Allgemeinen dazu tendieren, sich zu überschätzen, wenn es darum geht, ihre eigenen Fähigkeiten und Kompetenzen in verschiedenen sozialen und intellektuellen Gebieten zu bewerten (Dunning, Kruger, 1999).

Die beiden Autoren behaupten, dieses beobachtbare Verhalten komme daher, dass Menschen, welchen die Kompetenz in bestimmten Fähigkeiten fehle, dadurch auch jene Fähigkeit nicht besäßen, diese Inkompetenz zu erkennen. Bezeichnet wird dieser Umstand bei ihnen als „doppelte Belastung“ für die Personen. Sie definieren außerdem Inkompetenz in diesem Zusammenhang als „weniger kompetent als die peers“ (Dunning, Kruger, 1999), und nicht als generell inkompetent auf einem bestimmten Gebiet.

Überdies stützen sich Dunning und Kruger auf die Hypothese von Janet Metcalfe aus dem Jahre 1998, die besagt, dass Menschen von Natur aus eher jene Möglichkeiten auswählen, von denen sie glaubten, sie wären die optimalen und vernünftigsten (Metcalfe, 1998). Aufgrund dieser Beobachtung argumentieren sie, dass Menschen, welche nicht wissen können, dass sie falsch liegen, davon ausgehen, dass ihre Antworten die richtigen seien oder sie eben gut in Tests abgeschnitten hätten (Dunning, Kruger, 1999).

1.5.1 Die Studie zum Dunning-Kruger-Effekt

Die beiden Autoren führen insgesamt vier Studien zur Selbsteinschätzung durch, in denen sie Student*innen der Cornell-Universität zu ihren Kompetenzen hinsichtlich Humor (also die Kompetenz, Lustiges zu erkennen und feststellen zu können, was andere Menschen als lustig ansehen würden), logischen Denkens und englischer Grammatik befragt haben. Aufgrund der besagten Beobachtung, dass inkompetenten Personen anscheinend die

metakognitive Fertigkeit fehle, gute Leistungen im Allgemeinen sowie auch ihre eigene fehlende Leistung zu erkennen, stellen sie vier Hypothesen auf:

- Inkompetente Personen überschätzen ihre eigene Leistung und ihre eigenen Kompetenzen.
- Inkompetente Personen erkennen ihre eigenen Fähigkeiten und Leistungen sowie die Fähigkeiten und Leistungen anderer schlechter als kompetente Personen.
- Inkompetente Personen verkennen das wahre Niveau ihrer Kompetenz und Leistung beziehungsweise erkennen sie es schlechter als kompetente Personen.
- Inkompetente Personen können durch Üben ihrer Kompetenzen diese Einschätzung verbessern (Dunning, Kruger, 1999).

Um diese Hypothesen zu testen, führen sie die Studien nacheinander durch und bauen diese aufeinander auf, wohlweislich die Ergebnisse der vorhergehenden Studie miteinbeziehend. So können die Forscher, nachdem die Ergebnisse der ersten Studie ihre erste Hypothese unterstützt haben, anhand der zweiten Studie die Frage nach der zweiten Hypothese beantworten, ob die inkompetenten Studienteilnehmer*innen ihre eigenen Fähigkeiten überschätzt oder die der besseren, kompetenten Teilnehmer*innen unterschätzt haben. Die Ergebnisse weisen darauf hin, dass Ersteres der Fall sei. Interessant zu beobachten ist dabei, dass sich der Effekt am anderen Ende der Skala umkehrt: Je besser die Teilnehmer*innen in den Versuchen abschneiden, desto schlechter schätzen sie sich ein. So vermutet sich das beste Viertel der Teilnehmer*innen auf dem Niveau, welches tatsächlich dem zweitbesten Viertel entspricht. Generell fällt der geschätzte Durchschnitt aller Versuchsteilnehmer*innen in das tatsächliche Ergebnis des zweitbesten Viertels. Diese Teilnehmer*innen haben sich also durchschnittlich am besten eingeschätzt. Weiters ist auffallend, dass diese Verteilung bei allen vier Versuchen von Dunning und Kruger annähernd gleich ist.

Der dritten Hypothese, so erklären Dunning und Kruger, liegt nun ein Phänomen zugrunde, welches Lee Ross, David Greene und Pamela House im Jahre 1977 beschrieben haben. Diese drei Autor*innen beweisen anhand von vier Versuchen, dass Menschen ihre eigenen Verhaltensweisen, Entscheidungen und Überzeugungen als normal, gebräuchlich und

häufig einschätzen, während sie aber andere, divergierende Ansichten als selten oder sogar als unangemessen betrachten (Ross, Greene, House, 1977). Sie nennen diesen Effekt „*false-consensus*-Effekt“ (zu Deutsch „Effekt der falschen Übereinstimmung“) oder „*egocentric attribution*“, also etwa „egozentrische Zuweisung“ (Ross, Greene, House, 1977). Dunning und Kruger sehen nun in ihrer dritten Studie die dritte Hypothese bestätigt und können außerdem beweisen, dass die falsche Wahrnehmung der inkompetenten Personen von einem falschen Bild ihres Selbst ausgeht, die kompetenten Personen aber, also jene, welche sich unterschätzt haben, einem *false-consensus*-Effekt unterliegen. Sie gehen von einem falschen Bild von den anderen aus, nämlich, dass diese ähnlich gut abschneiden würden wie sie selbst (Dunning, Kruger, 1999).

Um ihre vierte Hypothese zu testen, bekommt die Hälfte der Studienteilnehmer*innen der vierten Studie nach dem ersten Durchgang ein kurzes Training das Themengebiet betreffend, um danach die vorher durchgeführten Tests noch einmal durchzusehen und einzuschätzen, ob und welche Items sie falsch beantwortet haben. Die andere Hälfte erhält kein solches Training. Tatsächlich verbessert sich bei den trainierten Teilnehmer*innen ihre Einschätzung, wie sie abgeschnitten hätten, bei den nicht trainierten hingegen verschlechtert sich diese sogar (Dunning, Kruger, 1999). Für die Autoren der Studie versteckt sich hier sogar ein Paradoxon: Um den inkompetenten Personen ihre Inkompetenz aufzuzeigen, müssten sie erst kompetent gemacht werden.

1.5.2 Erklärungen für den Dunning-Kruger-Effekt

Rückblickend auf die Ergebnisse der grundsteinlegenden Studie von 1999 und weiterer, aufbauender Studien, welche er später mit verschiedenen Autor*innen konzipiert hat (siehe unter anderem Dunning et al, 2003; Caputo, Dunning 2005; Carter, Dunning, 2008;), gibt David Dunning im Jahre 2011 mehrere Erklärungsmöglichkeiten zum Dunning-Kruger-Effekt (Dunning, 2011).

Zu Beginn spricht Dunning davon, dass die Unwissenheit der Personen ob ihres tatsächlichen Wissensstandes darauf zurückzuführen sei, dass sie nicht erkennen können, was sie nicht wissen. Das heißt, dass sie nicht einschätzen können, wie gut sie tatsächlich auf einem Gebiet sind, wenn sie die eigenen Unterlassungsfehler nicht erkennen können. Ein einfaches Beispiel ist folgendes: Personen werden aufgefordert,

Anwendungsmöglichkeiten des Schmieröles WD-40 aufzuzählen. Von diesen können theoretisch 50 oder mehr genannt werden, natürlich abhängig davon, wie viele Verwendungsmöglichkeiten eine Vergleichsperson oder etwa die Herstellerfirma angibt. Wenn nun jemand in einer Stunde „nur“ 15 davon anführen könnte, könnte diese Versuchsperson nicht einschätzen geschweige denn wissen, wie viele sie nicht aufgezählt hat oder wie ihre Leistung im Vergleich zu anderen Testpersonen ist, geschweige denn wissen, wie viele Anwendungen es insgesamt gibt. Dieses Beispiel wird von Deanna Caputo und David Dunning im Jahre 2005 im Zuge ihrer Untersuchungen angeführt, welche genau zum Thema des Wegfalles beziehungsweise Nicht-Wissens von Informationen bei Selbsteinschätzungen geforscht haben (Caputo, Dunning, 2005).

Dunning zufolge gibt es drei Hauptgründe dafür, warum dieses Nicht-Wissen nicht erkannt wird. Er bezeichnet diesen Umstand als „*meta-ignorance*“, zu Deutsch etwa „Meta-Unkenntnis“.

Erstens liege für viele Menschen ihr Nicht-Wissen in der Ebene des *unbekannten Unbekannten*, also Wissen, von dem sie nicht wissen, dass sie es nicht besitzen, im Gegensatz zu dem *bekannten Unbekannten* (Wissen, von dem diese Menschen wissen, dass sie es nicht besitzen) und natürlich dem *bekannten Bekannten*, Wissen, welches sie besitzen und dies ihnen bewusst ist (Dunning, 2011). Wenn also zum Beispiel eine Situation vorliegt, bei der eine Entscheidung getroffen oder etwa eine Frage beantwortet werden soll, kann eine Person, die auf dem jeweiligen Gebiet nicht besonders bewandert ist, nur auf bekannte Tatsachen zurückgreifen, also sowohl auf für diese Person Unbekanntes als auch Bekanntes, nicht jedoch auf unbekannte Tatsachen.

Der zweite Grund, dass das Nicht-Wissen nicht erkannt wird, sei ganz einfach der, dass das Wissen der Menschen einer Missinformation unterliegt (Dunning, 2011). Wenn Menschen falsche Informationen als richtig eingespeichert haben, ist es natürlich für sie quasi unmöglich, den wahren Stand der Dinge zu erkennen beziehungsweise einzusehen, dass sie ihr Wissen aus falschen Informationen beziehen. Dies kann so lange bestehen, bis sie jemand darauf hinweist, oder auch darüber hinaus, wenn die betroffene Person die neuen Informationen nicht annehmen will oder kann. Mary Kister Kaiser, John Jonides und Joanne Alexander untersuchen dieses Phänomen im Jahre 1986 anhand von Befragungen zu dem Wissen über physikalische Gesetzmäßigkeiten (Kister, Jonides, Alexander, 1986).

Als dritten Grund nennt David Dunning nun, sich beziehend auf Forschungen von Paulhus et al. (2003) sowie Bishop et al. (1986), dass Menschen auch auf Wissen zurückgreifen, das sie in bestimmten Gebieten haben, um damit Wissen zu kompensieren, welches ihnen in anderen Bereichen fehlt. Delroy Paulhus und seine Mitarbeiter*innen untersuchen dies im Jahre 2003, indem sie Menschen zu verschiedenen Gebieten befragen, unter die auch einige erfundene Themen gemischt sind (Paulhus, Harms, Bruce, Lysy, 2003). Die Versuchspersonen behaupten nun, selbst bei den Themen, welche es gar nicht gibt, ein gewisses Maß an Expertise zu besitzen, was natürlich nicht möglich sein kann. Dieser Effekt wird als „*Over-claiming*“ von Informationen benannt. Ein früherer Versuch dazu von George Bishop und seinen Kolleg*innen aus dem Jahre 1986 zeigt, dass dieser Effekt des *Over-claiming* noch verstärkt wird, wenn die Antwortmöglichkeit „Ich weiß nicht“ nicht zur Verfügung steht (Bishop, Tuchfarber, Oldendick, 1986). Menschen versuchen also, mit Informationen zu arbeiten, welche sie als relevant für die aktuelle Problemstellung betrachten, selbst wenn dieses Wissen nichts damit zu tun hat. David Dunning bezeichnet dies in seiner Erklärung auch als „*reach-around knowledge*“, also zu Deutsch etwa „ausweichendes Wissen“ (Dunning, 2011).

1.6 Hypothesen

Bezugnehmend auf die Error Management Theory von Haselton und Buss (Kapitel 1.1) und das Minimal Investment nach Trivers (Kapitel 1.2) soll erstens nachfolgend untersucht werden, ob die Beobachtung der evolutionär bedingten, kognitiven Eigenschaft der Selbstüberschätzung bei Männern beziehungsweise der Selbstunterschätzung bei Frauen auch in einem Kontext außerhalb des reproduktiven Hintergrundes zu machen ist. Dies geschieht in dieser Arbeit anhand zweier Geduldspiele (siehe Kapitel 2.5) und der Selbsteinschätzung der Versuchspersonen dazu. Demzufolge lautet die Hypothese zur EMT folgendermaßen:

Männer zeigen außerhalb des Reproduktionskontextes die Tendenz, sich optimistischer einzuschätzen, während Frauen dazu tendieren, sich zu unterschätzen.

Der zweite Schwerpunkt der Diplomarbeit wird darauf gelegt, einen möglichen Dunning-Kruger-Effekt nach Dunning und Kruger (Kapitel 1.5) und somit eine Illusory superiority (Kapitel 1.4) nach der Durchführung der Versuche mit den Geduldspielen zu verorten. Dabei wird vor allem auf die erste Erkenntnis von Dunning und Kruger eingegangen, welche besagt, dass Menschen, die keine oder wenig Erfahrung respektive Kompetenz in behandelten Gebieten haben, dies nicht wissen beziehungsweise nicht wissen können, und sich deswegen tendenziell überschätzen, während Menschen mit mehr Kompetenz dazu neigen, sich eher zu unterschätzen. Der besagte Kompetenzbereich liegt hierbei wiederum bei den Geduldspielen. Somit kann in Zusammenhang mit dem Dunning-Kruger-Effekt folgende, zweite Hypothese für die Arbeit aufgestellt werden:

Menschen, die weniger Erfahrung mit Geduldspielen haben, überschätzen sich in Korrelation mit dem Dunning-Kruger-Effekt tendenziell, während sich Menschen, welche viel Erfahrung mit Geduldspielen haben, tendenziell unterschätzen.

Es werden außerdem im Zuge der Befragungen weitere Items erörtert, wie etwa das Befinden der Studienteilnehmer*innen zum Zeitpunkt der Befragung als auch das geschätzte Können hinsichtlich Geduldspielen der Studienteilnehmer*innen. Daher werden bei den Ergebnissen zusätzlich die Fragen nach den Auswirkungen von Befinden und geschätztem Können auf die Selbsteinschätzung der Testpersonen beantwortet.

Nachdem die Fragestellungen als Grundlage für die Forschungsarbeit aufgestellt worden sind, sollen nun die verwendeten Methodiken aufgeführt sowie der Vorgang der praktischen Untersuchung dargelegt werden.

2. Methoden

2.1 Anzahl an Versuchsteilnehmer*innen

Für diesen Versuchsaufbau wurden in einem Zeitraum vom 30.10.2019 bis 20.11.2019 insgesamt 302 Testpersonen befragt, davon 153 weibliche und 149 männliche Testpersonen. Das Alter der weiblichen Teilnehmerinnen lag dabei zwischen 11 und 72 Jahren, das Alter der männlichen Teilnehmer zwischen 17 und 79 Jahren.

2.2 Auswahl Standorte

Die Befragungen wurden größtenteils in den Vormittags- bis Abendstunden an den folgenden Standorten durchgeführt: im Hauptgebäude der Universität Wien (Universitätsring 1, 1010 Wien), im Universitätsgebäude UZA 1, Althanstraße und im Universitätsgebäude UZA 2, Althanstraße im neunten Wiener Gemeindebezirk, im Rathauspark Wien (Friedrich-Schmitz-Platz 1, 1010 Wien) und im Innenhof des MuseumsQuartier Wien (Museumsplatz 1, 1070 Wien) sowie in einem Wohnhaus im 14. Wiener Gemeindebezirk. Grund für die Auswahl dieser verschiedenen, überwiegend öffentlichen Befragungsorte war die Gewährleistung einer bestmöglichen Erreichbarkeit der verschiedenen Altersgruppen.

2.3 Vorgehen bei der Datenerhebung

An den jeweiligen Standorten kam es nun zum Ansprechen von Personen, welche vorwiegend alleine anzutreffen waren, um eine Beeinflussung der Befragten durch eine/n mögliche/n Dritte/n zu unterbinden. Bei einer positiven Antwort der angesprochenen Personen folgte nun das Abarbeiten des Fragebogens mittels digitaler Online-Befragung auf einem Tablet und Ausgabe zweier Geduldspiele (im Fragebogen als „Probleme“ bezeichnet).

Während dieser Zeit der Befragung hatte nur der Versuchsleiter Einsicht in den Fragebogen, um den Versuchsteilnehmer*innen keine Vorwegnahme der einzelnen Items zu gewährleisten.

Der Grund für die Ausgabe von zwei verschiedenen Problemen liegt einerseits in der Möglichkeit, dadurch verschiedene Werte der Einschätzung anhand unterschiedlicher Geduldspiele zu ermitteln, andererseits stellt sich so dem Versuchsleiter die Frage nach der möglichen gegenseitigen Beeinflussung der beiden Probleme, je nachdem in welcher Reihenfolge diese ausgegeben und bearbeitet werden. Näheres zu diesen Überlegungen wird weiter unten bei den Ergebnissen in dieser Arbeit erläutert (siehe 3.3).

Es folgt eine nähere Beschreibung der bei den Befragungen verwendeten Mittel.

2.4 Fragebogen

Ausgefüllt wurde ein aus 27 Items bestehender Fragebogen, welcher vom Versuchsleiter eigens für den Versuchsaufbau entworfen worden war (s. Anhang). Realisiert wurde der Fragebogen mithilfe der online-Software „SurveyMonkey“ (WWW: <https://de.surveymonkey.com>).

Teilnehmer*innen wurden nach ihrem Alter, dem derzeitigen Befinden und nach einer Einschätzung der eigenen Erfahrung mit Geduldspielen und einer Einschätzung der eigenen Leistungen bei Geduldspielen befragt. Dabei sollten die Proband*innen ihr Befinden auf einer Skala von 1 („sehr schlecht“) bis 8 („sehr gut“) angeben. Die gleichen Skalen wurden bei der Einschätzung der eigenen Erfahrung und der eigenen Leistung bei Geduldspielen verwendet: von 1 („keine Erfahrung“) bis 8 („viel Erfahrung“) bei der Erfahrung respektive von 1 („sehr schlecht“) bis 8 („sehr gut“) bei der Leistung.

Des Weiteren wurden das Geschlecht und die sexuelle Orientierung erfragt sowie die Frage nach einer/m derzeitigen Lebenspartner*in gestellt.

War die befragte Person weiblich, wurde zusätzlich gefragt, ob die Menopause bei ihr bereits eingetreten sei. Entfiel die Antwort positiv (dies war bei 23 der 45 befragten Frauen über 40 Jahren der Fall, wobei die jüngste der Teilnehmerinnen 53 Jahre alt war), wurde die Teilnehmerin gefragt, ob sie Hormonersatzprodukte einnehme (nur 2 der 23 befragten Frauen bejahten dies). Entfiel sie negativ, wurden die Teilnehmerinnen nach der Dauer eines durchschnittlichen Menstruationszyklus und dem Eintreten der letzten Monatsblutung befragt. Außerdem wurde ermittelt, ob die Frauen hormonelle Verhütungsmittel einnahmen (36 aller 131 Frauen vor eingetretener Menopause bejahten dies). Dieser Teil entfiel bei den

Frauen, bei denen die Menopause bereits eingetreten war, im Fortlauf des Fragebogens automatisch.

Als zusätzliche Items wurden die Fragen nach einer derzeitigen Einnahme von Medikamenten sowie nach einem etwaigen Alkoholkonsum in den letzten 5 Stunden gestellt. Waren die Antworten positiv, so wurde dies in einem Kommentar vermerkt, sowohl die Art der Medikamente als auch die angegebene Menge an konsumiertem Alkohol.

Nach Erhebung dieser demografischen Daten kam es nun zum nächsten Schritt, dem Präsentieren des ersten Geduldspiels. Dabei wurde nach der Einschätzung der grundsätzlichen Lösbarkeit des Problems gefragt sowie danach, ob die Versuchspersonen glaubten, das vorgelegte Problem in zwei Minuten lösen zu können. Essenziell für die Studie war hierbei, dass die Proband*innen nicht wussten, dass es sich bei dem Thema der Studie um Selbsteinschätzung nach der EMT und dem Dunning-Kruger-Effekt handelt, um die Studienteilnehmer*innen nicht durch dieses Wissen in ihrer Einschätzung zu beeinflussen.

Danach wurde die Befragung für zwei Minuten unterbrochen und die Studienteilnehmer*innen versuchten sich an der Lösung der Probleme. Die Zeit wurde dabei mit der Timer-Funktion des Tablets genommen, nach spätestens zwei Minuten angehalten und somit der Versuch abgebrochen. Wurde das Spiel nicht gelöst, folgten einerseits die Frage, ob die Proband*innen glaubten, das Problem mit mehr Zeit lösen zu können, sowie andererseits eine neuerliche Frage nach der grundsätzlichen Lösbarkeit des Problems. Sollte das Problem von der Versuchsperson gelöst worden sein, wurde anschließend das zweite Geduldspiel ausgegeben, wobei der Ablauf jenem des ersten Problems glich. Die Abfolge der präsentierten Probleme wurde hierbei randomisiert, um die Beeinflussung auf die Selbsteinschätzung der Proband*innen und andere Störgrößen durch das jeweilige erste Problem zu minimieren und um statistisch sicherer arbeiten zu können.

Nach Beendigung beider Probleme und Beantwortung der dazugehörigen Fragen durch die Studienteilnehmer*innen wurde ihnen für die Teilnahme an diesem Experiment gedankt und somit die Befragung beendet.

2.5 Die Geduldspiele

Zur Datenerhebung dieses Versuches wurden zwei Geduldspiele unterschiedlicher Machart verwendet.

Bei dem ersten Spiel handelt es sich um das 3D-Puzzle „Metallpuzzle. Ring mit Stern“ der Firma Bartl GmbH, welches aus verchromtem Metall gefertigt ist. Es besteht aus einem elliptisch geformten Ring und einem Stern mit unterschiedlich langen Zacken, welcher von den Versuchsteilnehmer*innen ohne Anwendung von Hilfsmitteln oder Gewalt innerhalb der vorgegebenen Zeit von zwei Minuten aus dem Ring befreit werden muss. Dies gelang 46 von allen 302 Versuchsteilnehmer*innen, davon 16 Frauen und 30 Männern.

Das zweite Spiel, ebenfalls von der Firma Bartl GmbH, ist ein Legespiel namens „Die verflixten Spechte“ und besteht aus neun verschieden bedruckten Holzplatten, auf denen sich jeweils zwei Oberkörper mit Köpfen und zwei Unterkörper verschiedenfarbiger Spechte befinden, welche sich, wenn richtig aneinandergefügt, zu einem Bild vereinen.

Hierbei muss von den Testpersonen genauestens beobachtet werden, ob die Ober- und Unterteile farblich und dem Körperbau entsprechend zusammenpassen. Dieses Problem bleibt nach außen hin offen, es gibt also keine passenden Gegenstücke an den Außenrändern, außerdem ergibt das Problem nur in einer quadratischen Form das richtige Endergebnis.

Dieses Problem wurde von insgesamt fünf der 302 Versuchsteilnehmer*innen gelöst, davon war eine weibliche Teilnehmerin und vier männliche Teilnehmer.

2.6 Auswertung

Die statistische Auswertung der ermittelten Werte und Daten erfolgte mithilfe des Statistik-Programmes SPSS in der Version 26.

Die erste Hypothese, nämlich die Frage nach den geschlechtsspezifischen Einschätzungen im Zuge der Error Management Theory wurde mittels dem Chi-Quadrat-Test (χ^2) nach Pearson ausgewertet. Um die Auswirkungen des Dunning-Kruger-Effektes zu observieren, wurde die zweite Hypothese mit Hilfe eines t-Tests überprüft.

3. Ergebnisse

3.1 Selbsteinschätzung nach der Error Management Theory

Die Hypothese hierzu lautet, dass Männer im nicht-reproduktiven Kontext zu einer optimistischeren Selbsteinschätzung tendieren, sich also tendenziell überschätzen, während Frauen eher zu einer pessimistischeren Selbsteinschätzung tendieren, sich also eher unterschätzen. Um die Übersicht über die Ergebnisse zu erleichtern, wurden die Ergebnisse den Geduldspielen nach geordnet.

3.1.1 Puzzlespiel „Ring mit Stern“

Vor der Manipulation mit dem Geduldspiel

Bei der Frage nach der grundsätzlichen Lösbarkeit des Geduldspiels „Ring mit Stern“ konnte geschlechtsspezifisch kein auffallender Unterschied festgestellt werden ($\chi^2 = 1,845$, $p = 0,174$). Bei den Männern beantworteten die Frage 18 mit „Nein“, 131 mit „Ja“, bei den Frauen waren es 27 mit „Nein“ und 126 mit „Ja“.

Ein signifikanter Unterschied wurde jedoch bei der Frage „Glauben Sie, dass Sie dieses Problem in zwei Minuten lösen können?“ beobachtet ($\chi^2 = 3,971$, $p = 0,046$): Hier antworteten 91 Männer mit „Nein“ und 58 Männer mit „Ja“, während bei den Frauen 110 mit „Nein“ und 43 mit „Ja“ antworteten. Männer schätzten sich hier also etwas optimistischer ein als Frauen (Abbildung 3).

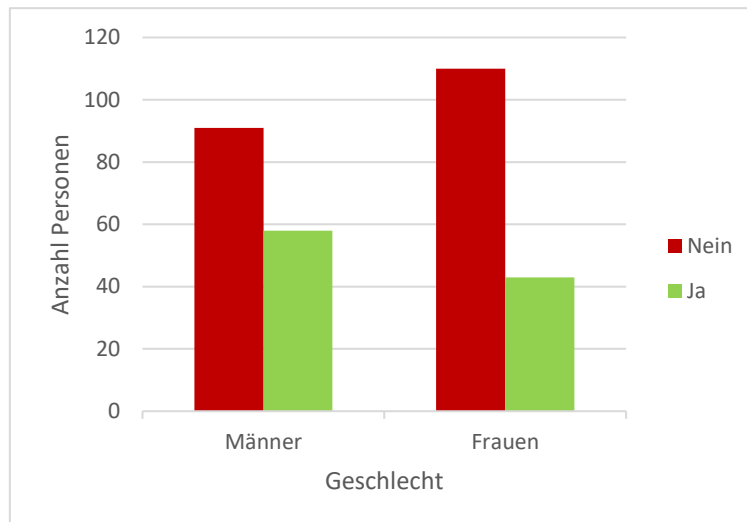


Abbildung 1: Selbsteinschätzung vor der Manipulation, $p = 0,046$

Nach der Manipulation mit dem Geduldspiel

Bei Studienteilnehmer*innen, welche das Problem nicht lösen konnten ($N_{\text{Gesamt}} = 257$), wurde ebenfalls kein geschlechtsspezifischer Unterschied bei der Frage gefunden, ob das Geduldspiel „Ring mit Stern“ grundsätzlich mit mehr Zeit schaffbar sei ($\chi^2 = 1,566$, $p = 0,211$). 36 Männer beantworteten diese Frage mit „Nein“, 97 Männer mit „Ja“, bei den Frauen waren es 48 negierte Antworten und 92 bejahte Antworten.

Jedoch erneut einen, diesmal sehr signifikanten Unterschied ergab nach erfolgloser Manipulation die Frage nach der Selbsteinschätzung, ob die Teilnehmer*innen glaubten, mit mehr Zeit das Problem lösen zu können ($\chi^2 = 9,631$, $p = 0,002$). Wieder schätzten sich Männer hier optimistischer ein als Frauen: 65 Männer beantworteten dies mit „Nein“ und 54 mit „Ja“, während 101 Frauen dies mit „Nein“ und 37 mit „Ja“ beantworteten (s. Abbildung 4).

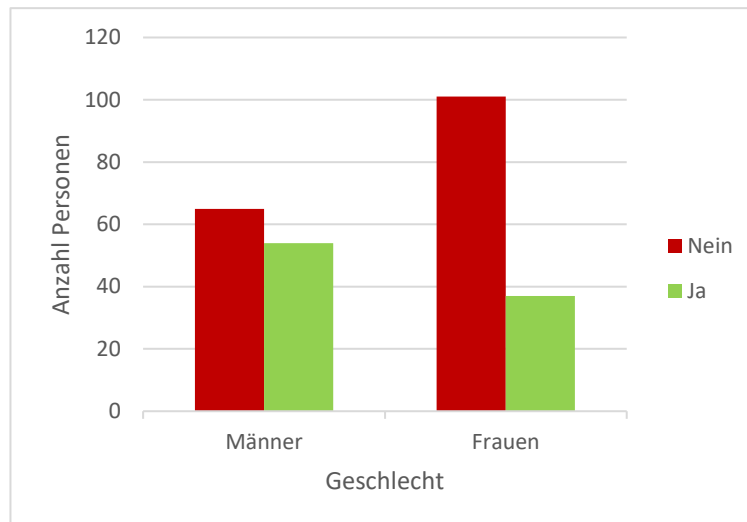


Abbildung 2: Selbsteinschätzung nach der Manipulation, $p = 0,002$

3.1.2 Legespiel „Die verflixten Spechte“

Vor der Manipulation mit dem Geduldspiel

Die Frage, ob die Versuchspersonen glaubten, das Problem sei grundsätzlich lösbar, beantworteten 13 Männer mit „Nein“ und 136 Männer mit „Ja“, beziehungsweise 13 Frauen mit „Nein“ und 140 Frauen mit „Ja“. Es gibt hier also keinen signifikanten Unterschied zwischen den Geschlechtern ($\chi^2 = 0,005$, $p = 0,944$).

Ebenso wenig signifikant unterschieden sich die Selbsteinschätzungen bei der Frage, ob die Proband*innen glaubten, das Problem in zwei Minuten lösen zu können ($\chi^2 = 0,232$, $p = 0,630$). Hier verneinten 83 Männer die Frage und 66 Männer bejahten sie, während 81 Frauen mit „Nein“ antworteten und 72 Frauen mit „Ja“.

Nach der Manipulation mit dem Geduldspiel

Insgesamt konnten $N_{\text{Gesamt}} = 297$ Versuchsteilnehmer*innen das Problem nicht in zwei Minuten lösen. Die Frage, ob sie glaubten, das Problem sei mit mehr Zeit grundsätzlich zu lösen, beantworteten 30 Männer mit „Nein“ und 115 Männer mit „Ja“; bei den Frauen waren es 33 „Nein“-Antworten und 119 „Ja“-Antworten. Es gibt hier also keinen statistisch signifikanten Unterschied zwischen den Geschlechtern ($\chi^2 = 0,046$, $p = 0,830$).

Die Selbsteinschätzung nach der Manipulation zeigt ebenso keinen signifikanten Unterschied ($\chi^2 = 0,632$, $p = 0,427$). Wurden die Studienteilnehmer*innen danach gefragt, ob sie glaubten, das Problem mit mehr Zeit lösen zu können, so antworteten 58 Männer mit „Nein“ und 87 Männer mit „Ja“, während 54 Frauen die Frage mit „Nein“ und 98 Frauen mit „Ja“ beantworteten.

3.2 Selbsteinschätzung nach dem Dunning-Kruger-Effekt

Die Hypothese zum Dunning-Kruger-Effekt lautet in dieser Diplomarbeit, dass sich Menschen mit mehr Erfahrung mit Geduldspielen tendenziell unterschätzen, während Menschen, die weniger Erfahrung auf dem Gebiet haben, sich eher überschätzen. Des Weiteren wurden auch die Auswirkungen zweier anderer Faktoren auf die Selbsteinschätzung mit in Betracht gezogen, namentlich das Befinden der Teilnehmer*innen und die eigene Leistung, nachfolgend mit „Können“ bezeichnet. Bei allen drei Sachverhalten wird, dem t-Test zufolge, der Mittelwert M_x angegeben. Das „x“ wird hierbei für die respektiven Faktoren stehen.

Auch hier werden die Ergebnisse der Übersicht halber den Geduldspielen nach geordnet, außerdem werden sowohl die Gesamtergebnisse beider Geschlechter als auch die geschlechtsspezifischen Ergebnisse präsentiert.

3.2.1 Puzzlespiel „Ring mit Stern“

Gesamtergebnisse

Die Anzahl der Versuchsteilnehmer*innen betrug bei diesem Versuch $N_{\text{Gesamt}} = 301$. Davon beantworteten 200 Teilnehmer*innen die Frage, ob sie das Problem in zwei Minuten lösen könnten, mit „Nein“ und 101 Teilnehmer*innen mit „Ja“.

Auswirkung der Erfahrung und des Könnens auf die Selbsteinschätzung

Es konnte gezeigt werden, dass die Erfahrung der Teilnehmer*innen mit Geduldspielen keine statistisch relevante Auswirkung auf die Selbsteinschätzung hatte. Bei einem

Mittelwert von $M_{\text{Erf.}} = 3,57$ bei den 200 „Nein“-Antworten und einem Mittelwert von $M_{\text{Erf.}} = 3,71$ bei den 101 „Ja“-Antworten betrug $t = -0,849$ ($p = 0,396$).

Das Können der Studienteilnehmer*innen zeigte ebenfalls keine relevante Auswirkung auf die Selbsteinschätzung ($t = -1,744$, $p = 0,082$). Die Mittelwerte betrugen bei 200 „Nein“-Antworten $M_{\text{Kön.}} = 3,31$ beziehungsweise $M_{\text{Kön.}} = 3,62$ bei 101 „Ja“-Antworten.

Auswirkung des Befindens auf die Selbsteinschätzung

Die Mittelwerte der gesamten Teilnehmer*innen betrugen hierbei für 200 „Nein“-Antworten $M_{\text{Bef.}} = 6,16$ und für 101 „Ja“-Antworten $M_{\text{Bef.}} = 6,15$. Es konnte auch hier keinerlei Auswirkung auf die Selbsteinschätzung festgestellt werden ($t = 0,097$, $p = 0,923$).

Beim Geduldspiel „Ring mit Stern“ kann also festgehalten werden, dass weder Erfahrung noch Können noch das Befinden der Teilnehmer*innen eine Auswirkung auf die Selbsteinschätzung hatten.

Ergebnisse nach Geschlecht – Männer

Es konnten bei den Männern keine statistisch relevanten Ergebnisse beobachtet werden. Von 149 Männern antworteten 91 mit „Nein“ und 58 mit „Ja“ auf die Frage, ob sie glaubten, das Problem in zwei Minuten lösen zu können.

Weder die Erfahrung ($t = -1,117$, $p = 0,266$ bei Mittelwerten von $M_{\text{Erf.}} = 3,51$ für die 91 „Nein“-Antworten und $M_{\text{Erf.}} = 3,76$ für die 58 „Ja“-Antworten), noch das Können ($t = -1,233$, $p = 0,220$ bei Mittelwerten von $M_{\text{Kön.}} = 3,29$ für die 91 „Nein“-Antworten und $M_{\text{Kön.}} = 3,57$ für die 58 „Ja“-Antworten), noch das Befinden ($t = 0,606$, $p = 0,545$ bei Mittelwerten von $M_{\text{Bef.}} = 6,16$ für die 91 „Nein“-Antworten und $M_{\text{Bef.}} = 6,07$ für die 58 „Ja“-Antworten) zeigten eine relevante Auswirkung auf die Selbsteinschätzung der männlichen Probanden.

Ergebnisse nach Geschlecht – Frauen

Selbige Ergebnisse zeigten die Auswertungen der Daten bei den Frauen. 152 Frauen versuchten sich an dem Geduldspiel „Ring mit Stern“, 109 davon beantworteten die Frage nach der Lösbarkeit des Problems in zwei Minuten mit „Nein“, 43 mit „Ja“.

Die drei untersuchten Items zeigten keine relevante Auswirkung auf die Selbsteinschätzung: Bei der Erfahrung wurden Mittelwerte von $M_{\text{Erf.}} = 3,62$ für die 109 „Nein“-Antworten beziehungsweise $M_{\text{Erf.}} = 3,65$ für die 43 „Ja“-Antworten ermittelt ($t = -0,107$, $p = 0,915$), für das Können Mittelwerte von $M_{\text{Kön.}} = 3,32$ für die 109 negativen Antworten und $M_{\text{Kön.}} = 3,70$ für die 43 positiven Antworten ($t = -1,289$, $p = 0,199$). Beim eigenen Befinden zeigten sich Mittelwerte von $M_{\text{Bef.}} = 6,16$ für die 109 „Nein“-Antworten und $M_{\text{Bef.}} = 6,26$ für die 43 „Ja“-Antworten ($t = -0,550$, $p = 0,583$).

3.2.2 Legespiel „Die verflixten Spechte“

Gesamtergebnisse

Die Gesamtanzahl an Versuchsteilnehmer*innen betrug bei diesem Versuch $N_{\text{Gesamt}} = 301$. Die Frage, ob sie das Problem in zwei Minuten lösen könnten, beantworteten 163 davon mit „Nein“ und 138 davon mit „Ja“.

Auswirkung der Erfahrung und des Könnens auf die Selbsteinschätzung

Bei einem Mittelwert von $M_{\text{Erf.}} = 3,64$ bei den 163 „Nein“-Antworten und einem Mittelwert von $M_{\text{Erf.}} = 3,59$ bei den 138 „Ja“-Antworten konnte keine relevante Auswirkung der Erfahrung gezeigt werden ($t = 0,275$, $p = 0,784$). Ebenso verhält es sich mit dem Können ($M_{\text{Kön.}} = 3,38$ bei „Nein“, $M_{\text{Kön.}} = 3,45$ bei „Ja“, $t = -0,396$, $p = 0,692$). Das bedeutet, dass aus den Ergebnissen dieses Versuches nicht ersichtlich wird, ob mehr Erfahrung oder ein besseres Können der Versuchsteilnehmer*innen tatsächlich ein Unterschätzen mit sich bringt.

Auswirkung des Befindens auf die Selbsteinschätzung

Anders verhält es sich hinsichtlich der Selbsteinschätzung unter Berücksichtigung des Befindens der Proband*innen. Es zeigt sich, dass sich Menschen mit einem besseren Befinden sehr signifikant pessimistischer einschätzen ($t = 2,834$, $p = 0,005$) bei einem Mittelwert von $M_{\text{Bef.}} = 6,30$ bei 163 „Nein“-Antworten im Gegensatz zu einem Mittelwert von $M_{\text{Bef.}} = 5,99$ bei 183 „Ja“-Antworten.

Diesen Ergebnissen lässt sich somit resümierend entnehmen, dass beim Geduldspiel „Die verflixten Spechte“ geschlechtsunspezifisch weder mehr Erfahrung noch besseres Können einen Einfluss auf die Selbsteinschätzung haben, also weder ein Über- noch ein Unterschätzen ersichtlich werden, jedoch Menschen, deren Befinden besser ist, dazu tendieren, sich zu unterschätzen.

Ergebnisse nach Geschlecht – Männer

Betrachtet man nun die Ergebnisse, welche Auswirkungen die Faktoren Erfahrung, Können und Befinden bei Männern auf die Selbsteinschätzung haben, können nachfolgende Beobachtungen gemacht werden. Die Anzahl an männlichen Teilnehmern war $N_{\text{Gesamt}} = 149$. Davon beantworteten 83 die Frage nach der Lösbarkeit des Problems in zwei Minuten mit „Nein“ und 66 mit „Ja“.

Auswirkung der Erfahrung und des Könnens auf die Selbsteinschätzung

Wie auch bei den Gesamtergebnissen zeigten weder persönliche Erfahrung der männlichen Probanden noch ihr Können statistisch signifikante Auswirkungen. Bei der Erfahrung verhält sich ein Mittelwert von $M_{\text{Erf.}} = 3,55$ für „Nein“-Antworten zu einem Mittelwert von $M_{\text{Erf.}} = 3,67$ für „Ja“-Antworten ($t = -0,504$, $p = 0,615$), beim Können beträgt $t = -0,705$ und $p = 0,482$ bei Mittelwerten von $M_{\text{Kön.}} = 3,33$ für 83 x „Nein“ und einem Mittelwert von $M_{\text{Kön.}} = 3,48$ bei 66 x „Ja“.

Auswirkung des Befindens auf die Selbsteinschätzung

Als korrelierend mit den Gesamtergebnissen, wenn auch nicht ganz so signifikant ($t = 2,208$, $p = 0,029$) erwies sich die Auswirkung des eigenen Befindens auf die eigene Einschätzung der männlichen Teilnehmer. Der Mittelwert von $M_{\text{Bef.}} = 6,28$ der 83 „Nein“-Antworten im Gegensatz zum Mittelwert von $M_{\text{Bef.}} = 5,94$ der 66 „Ja“-Antworten zeigt, dass auch hier ein besseres Befinden der männlichen Probanden zu einer pessimistischeren Selbsteinschätzung führt.

Zusammenfassend kann gesagt werden, dass bei den männlichen Versuchsteilnehmern zwar ihre Erfahrung und ihr Können keine Auswirkung auf ihre Selbsteinschätzung zeigen, das Befinden hingegen insofern einen Effekt zeigt, als sich jene Männer eher unterschätzen, deren Befinden durchschnittlich besser ist.

Ergebnisse nach Geschlecht – Frauen

Insgesamt nahmen $N_{\text{Gesamt}} = 152$ Frauen an diesem Versuch teil, die Frage nach der Lösbarkeit des Problems ins zwei Minuten beantworteten davon 80 mit „Nein“ und 72 mit „Ja“.

Auswirkungen der Erfahrung und des Könnens auf die Selbsteinschätzung

Es konnte kein Zusammenhang zwischen Erfahrung oder Können und der Selbsteinschätzung der Versuchsteilnehmerinnen gefunden werden ($t = 0,862$, $p = 0,390$ bei einem Mittelwert von $M_{\text{Erf.}} = 3,73$ der „Nein“-Antworten und einem Mittelwert von $M_{\text{Erf.}} = 3,53$ der „Ja“-Antworten bei der Erfahrung sowie einem $t = 0,079$, $p = 0,937$ bei einem Mittelwert von $M_{\text{Kön.}} = 3,44$ der „Nein“-Antworten und einem Mittelwert von $M_{\text{Kön.}} = 3,42$ der „Ja“-Antworten beim Können). Auch hier hat also die Erfahrung mit Geduldspielen und das Können keinerlei Auswirkung auf die Selbsteinschätzung.

Auswirkung des Befindens auf die Selbsteinschätzung

Bei diesem Item findet sich bei den Frauen mit $t = 1,833$ und $p = 0,069$ eine Tendenz in Richtung pessimistischerer Einschätzung bei besserem Befinden, welche mit den Gesamtergebnissen korreliert, jedoch nicht relevant ist. Die Mittelwerte waren hierbei $M_{\text{Bef.}} = 6,33$ für die 80 „Nein“-Antworten und $M_{\text{Bef.}} = 6,03$ für die 72 „Ja“-Antworten.

Zusammenfassend lässt sich sagen, dass, einhergehend mit den Gesamtergebnissen, das bessere Befinden der Frauen tendenziell in Richtung Unterschätzung führt, während wiederum weder die Erfahrung noch das Können eine statistisch signifikante Auswirkung auf die Selbsteinschätzung der Frauen haben.

3.3 Weitere interessante Ergebnisse

Nennenswert, da interessant zu beobachten, war folgender Punkt: Statistisch gesehen zeigte sich eine Auswirkung auf die Einschätzung NACH der Manipulation durch das zweite Problem, wenn das erste Problem erfolgreich gelöst worden war. Menschen tendierten dazu, sich dann optimistischer einzuschätzen. Dieses Phänomen konnte sowohl bei der Frage, ob die Versuchspersonen glaubten, das Problem mit mehr Zeit lösen zu können, als auch bei der Frage, ob das Problem grundsätzlich nach erfolgloser Manipulation lösbar sei, beobachtet werden. Der verwendete Test war, wie auch schon bei der Auswertung der Ergebnisse zur Error Management Theory, wieder ein Chi-Quadrat-Test nach Pearson. Vor der Manipulation durch das zweite Problem ergaben sich für $N_{\text{Gesamt}} = 279$ Testteilnehmer*innen folgende Werte: 170 Teilnehmer*innen glaubten, sie könnten das Problem nicht in zwei Minuten lösen, 109 Teilnehmer*innen bejahten diese Frage ($\chi^2 = 0,035$, $p = 0,851$). Hier wird also keine Relevanz zur Einschätzung deutlich. Nach der Manipulation zeigte sich zur Frage, ob die Proband*innen glaubten, das Problem mit mehr Zeit lösen zu können, eine hoch signifikante Auswirkung hin zu einer optimistischen Einschätzung mit 128 „Nein“-Antworten und 151 „Ja“-Antworten ($N_{\text{Gesamt}} = 279$, $\chi^2 = 17,818$, $p = <0,001$). Ebenso signifikant zugunsten einer optimistischeren Einschätzung fiel das Ergebnis zur Frage nach der grundsätzlichen Lösbarkeit des Problems aus ($\chi^2 = 5,931$, p

= 0,015): Nur 62 Versuchsteilnehmer*innen beantworteten diese Frage mit „Nein“, jedoch 217 Teilnehmer*innen mit „Ja“ ($N_{\text{Gesamt}} = 279$).

Ein weiterer beachtenswerter Aspekt ist, dass augenscheinlich die Reihenfolge der Ausgabe der Probleme keinerlei Auswirkung auf die Selbsteinschätzung der Studienteilnehmer*innen hatte. Wurde zum Beispiel das Legespiel „Die verflixten Spechte“ zuerst ausgegeben, hatte dies keine statistisch relevante Auswirkung auf die Einschätzung des darauffolgend ausgegebenen Puzzlespiels „Ring mit Stern“. Von $N_{\text{Gesamt}} = 252$ befragten Personen antworteten 165 mit „Nein“ und 87 mit „Ja“, als sie gefragt wurden, ob sie glaubten, das Problem „Ring mit Stern“ mit mehr Zeit lösen zu können ($\chi^2 = 0,003$, $p = 0,960$). Die Frage, ob sie glaubten, das Problem sei grundsätzlich lösbar, beantworteten nach erfolgloser Manipulation 81 Versuchsteilnehmer*innen mit „Nein“ und 171 mit „Ja“ ($\chi^2 = 0,594$, $p = 0,441$).

4. Diskussion

4.1 Interpretation der Selbsteinschätzung nach der EMT

Die Ergebnisse zum Geduldspiel „Ring mit Stern“ (siehe 3.1.1) zeigen, dass sich Männer bei diesem Beispiel tendenziell optimistischer einschätzen als Frauen, wenn sie schätzen sollen, ob das Problem in zwei Minuten lösbar ist. Dieser Effekt ist sowohl vor der Manipulation durch das Geduldspiel ($p = 0,046$) als auch nach der (erfolglosen) Manipulation durch das Geduldspiel, hier sogar hoch signifikant ($p = 0,002$) zu beobachten. So kann die erste Hypothese in Bezug auf das Problem „Ring mit Stern“, nämlich dass nach der Error Management Theory von Haselton und Buss Männer dazu neigen, sich zu überschätzen (Haselton, Buss, 2000), teilweise bestätigt werden. Dies unterstreicht die bisherigen Forschungsergebnisse der Literatur (Galperin, Haselton, 2013; Haselton, Buss, 2000), welche in Kapitel 1.3 zusammengefasst sind. Bezogen auf die zweite Hälfte der Hypothese kann bei diesem Versuch nicht statistisch relevant gezeigt werden, dass sich Frauen tendenziell unterschätzen. Ein Grund hierfür könnte ein vorliegender Versuchsleitereffekt sein. Der Umstand, dass der Versuch zu dieser Diplomarbeit ausschließlich von einem Mann durchgeführt worden ist, legt die Annahme nahe, dass eine unbewusste Beeinflussung der Testteilnehmer*innen stattgefunden hat. Theorien dazu könnten sein, dass Frauen sich angesichts des männlichen Versuchsleiters eher positiver einschätzen würden, da sie ihm imponieren möchten. Dies könnte eine weitere Erklärung dafür sein, dass die in dieser Diplomarbeit erhobenen Daten bezüglich der Error Management Theory die Hypothese zur Unterschätzung der Frauen nicht unterstützen. Umgekehrt wäre eine interessante Frage, wie sich die Selbsteinschätzung von Männern verändert, wenn die Untersuchung von einer Frau oder einem Mann geleitet wird. Überlegungen hierzu können sein, dass auch Männer einem männlichen Versuchsleiter gegenüber eine positivere Selbsteinschätzung zeigen, um die eigenen Stärken und Vorteile hervorzuheben. Weiter unten wird im Ausblick auf künftige Studien näher beschrieben, wie diese Überlegungen ferner erforscht werden können (Kapitel 4.3).

Es soll hier weiters erwähnt werden, dass die Auswahl der Befragungsorte eine Auswirkung auf die Einschätzung der befragten Personen gehabt haben könnte. Der beträchtlicher Teil der Befragungen fand im Hauptgebäude der Universität Wien statt, vor allem in der Nähe des Student Space und den umliegenden Seminarräumen. Es fiel auf,

dass die dort gehaltenen Seminare und Vorlesungen, vorwiegend aus dem geisteswissenschaftlichen Bereich, von überdurchschnittlich vielen Frauen besucht wurden. Es kann daher sein, dass Frauen, welche sich im Umfeld von anderen Frauen wähnen, durch diesen Umstand selbstbewusster auftreten und mehr auf ihre eigene Meinung und ihr Können vertrauen. Eine Gegenstudie an anderen Standorten mit Studienrichtungen, welche männerlastiger sind, könnte hierzu Vergleichsdaten liefern.

Beim Geduldspiel „Die verflixten Spechte“ zeigen die Resultate keinerlei relevante Auswirkungen der verschiedenen Items auf die Selbsteinschätzung von Frauen und Männern, weder vor noch nach der Manipulation durch das Geduldspiel. Dies könnte auf die grundsätzlich geringe Anzahl von $N = 302$ an befragten Personen zurückzuführen sein. Weiters sind andere Faktoren, welche im Zuge der Befragung erhoben worden sind, nicht in die statistische Auswertung miteinbezogen. Zu diesen gehören zum Beispiel der Einfluss von Alkohol auf die Teilnehmer*innen, der unterschiedliche Hormonspiegel bei Frauen aufgrund ihrer Monatszyklusphase oder die sexuelle Orientierung der Teilnehmer*innen. Alle diese Faktoren könnten eine Auswirkung zeigen, sind aber bei der Auswertung im Rahmen dieser Abschlussarbeit nicht berücksichtigt worden, da dies zu umfangreich wäre.

Eine weitere aufschlussreiche Überlegung zu dem Gebiet könnte unter anderem sein, welche Einflüsse heutzutage die gesellschaftlichen Aspekte auf die Einschätzung von Männern und Frauen haben. Vor allem im reproduktiven Kontext könnte man darüber diskutieren, ob sich die soziale Entwicklung, auch mit Hinsicht auf hormonelle und mechanische oder operative Verhütungsmethoden, in Richtung einer Annullierung oder zumindest Abschwächung von Error Management und Minimal Investment Theorien verschiebt.

4.2 Interpretation der Selbsteinschätzung nach Dunning-Kruger

Die Hypothese, welche in Kapitel 1.6 aufgestellt ist, besagt, dass sich Menschen nach dem Dunning-Kruger-Effekt eher unterschätzen, wenn sie mehr Erfahrung auf einem Gebiet haben, und sich umgekehrt dann überschätzen, je weniger Erfahrung sie haben. Diese Aussagen konnten im Zuge der beiden Geduldspiele und den dabei erhobenen Daten nicht gestützt werden. Weder hinsichtlich des Puzzlespiels „Ring mit Stern“ noch beim Legespiel „Die verflixten Spechte“ kann eine statistisch relevante Auswirkung der Erfahrung oder des

Könnens auf die Selbsteinschätzung nachgewiesen werden. Dies kann unterschiedliche Hintergründe haben: Ein Argument könnte wiederum die geringe Anzahl von $N = 302$ Versuchsteilnehmer*innen sein. Ein größerer Datensatz kann natürlich genauere Ergebnisse liefern. Ein weiterer, wichtiger Punkt ist die sehr subjektive Einschätzung der Testpersonen von Kompetenzlevel, sprich Erfahrung und Können, auf dem Gebiet der Geduldspiele. Dies liegt aber offensichtlich in der Natur der Sache. Der Versuchsleiter kann beim Auswerten der Datensätze nur auf die subjektiv richtige Einschätzung der Proband*innen vertrauen.

Interessanterweise zeigt sich jedoch eine Auswirkung des Befindens der Proband*innen auf ihre Einschätzung zu Geduldspielen. Betrachtet man die Auswertung der Daten des Legespiels „Die verflixten Spechte“, so ist erkennbar, dass sich Proband*innen, welche ein besseres Befinden aufweisen, eher pessimistischer einschätzen, und zwar sehr deutlich signifikant ($p = 0,005$). Untersucht man diesen Umstand getrennt nach Geschlechtern, zeigt sich, dass der Effekt bei Männern stärker ($p = 0,029$) als bei Frauen ausgeprägt ist, welche aber eine Tendenz dazu zeigen ($p = 0,069$). Diese Erkenntnis scheint etwas überraschend, möchte man doch im ersten Moment meinen, Menschen mit einer besseren Stimmung würden eher dazu tendieren, sich zu überschätzen und vice versa. Es können die Gründe für dieses Phänomen mannigfaltig sein. Einer davon könnte sein, dass sich Menschen, die ein besseres Befinden haben, als eine Art Selbstschutzmechanismus eher pessimistischer einschätzen. Sollte sich ihr tatsächliches Ergebnis als so schlecht oder schlechter erweisen, würden sie sich damit möglicherweise ihre gute Laune nicht selbst vermiesen.

Bemerkenswert ist bei den Ergebnissen zum Dunning-Kruger-Effekt jene Erkenntnis, dass die Reihenfolge der Geduldspiele keine Auswirkung auf die Einschätzung des Problems zeigt, welches als zweites ausgegeben wird ($p = 0,960$ für die Selbsteinschätzung, $p = 0,441$ für die grundsätzliche Lösbarkeit). Eigentlich ist die Erwartungshaltung dahingehend anders gewesen, namentlich, dass die Einschätzungen sich ins Positive oder Negative verändern würden, je nachdem, ob das erste Problem als schwer oder leicht angesehen worden ist. Da der Schwerpunkt dieser Diplomarbeit auf anderen Parametern liegt, wird hier keine weitere Interpretation vorgenommen, wenngleich die vorliegenden Daten weitere Interpretations- und Vergleichsmöglichkeiten bieten würden.

4.3 Ausblick auf künftige Studien

Es sind im Rahmen der Befragungen noch einige weitere Items respektive Einflussfaktoren von den Proband*innen erfragt worden, deren mögliche Auswirkungen auf die Selbsteinschätzung in dieser Diplomarbeit auf Grund deren zu geringer Relevanz für diese Arbeit nicht untersucht werden. Zu diesen Faktoren gehören der Einfluss von Alkohol und Medikamenten auf die Personen, die sexuelle Orientierung und die Frage nach einer/m Lebenspartner*in. Frauen sind zusätzlich nach der Dauer ihres Monatszyklus gefragt worden und nach dem Datum ihrer letzten Menstruation, ob sie hormonell verhüten würden und ob sie bereits ihre Menopause gehabt hätten. Der unterschiedlich hohe Hormonhaushalt bei Frauen könnte interessante Auswirkungen auf die Selbsteinschätzung haben. Es könnte sich zum Beispiel die Frage gestellt werden, ob sich Frauen, welche hormonell verhüten und dadurch eine erhöhte Dosis an weiblichen Hormonen im Körper haben, sich eher über- oder eher unterschätzen würden. Man könnte annehmen, dass die größere Menge an weiblichen Hormonen das Selbstbewusstsein und damit auch das Befinden und die Selbsteinschätzung der Frauen beeinflussen könnte. Daher sind beide Theorien plausibel. Weiters kann auch argumentiert werden, dass zum Zeitpunkt des Eisprungs sowie vor und nach der Regelblutung die Konstellation und Menge an Hormonen variieren und somit auch einen Einfluss haben könnten.

Weiters könnte zusätzlich untersucht werden, ob eventuell ein Versuchsleitereffekt vorliegt. Es könnte zum Beispiel eine Metastudie angelegt werden mit den Daten aus dieser Diplomarbeit und vorherigen Arbeiten zu dem Thema, bei denen Frauen den Versuch geleitet haben. Alle diese Faktoren können als Vorlage zu weiteren Forschungen auf dem Gebiet der Error Management Theory dienen, welches noch lange nicht ausgereizt ist.

Interessant wäre auch zu beobachten, wie sich die Menge an Bewegung und ausgeübtem Sport auf die Selbsteinschätzung bei Menschen auswirkt. Studien zeigen, dass mehr Sport sich in einem gesteigerten Selbstbewusstsein und einem gesteigerten Selbstwertgefühl äußern kann (Kubesch, 2002; Opdenacker, Delecluse, Boen, 2009; Sonstroem, Morgan, 1989). Ein besseres Selbstbewusstsein könnte nun wiederum zu einer besser Selbsteinschätzung in verschiedenen Domänen führen. Hierfür könnte eine Gegenstudie zum Beispiel an der USI Wien durchgeführt werden.

Auch zum Dunning-Kruger-Effekt können viele weitere Nachforschungen betrieben werden, zumal die Selbsteinschätzung bei Menschen auf so vielen verschiedenen Gebieten unterschiedlich ist. Zum einen können zu einer genaueren Erhebung eine viel größere Zahl an Befragungen durchgeführt werden, zum anderen könnte in eigenen Arbeiten zuerst das wahre Kompetenzniveau der Teilnehmer*innen erhoben werden. Dies kann zu noch genaueren Einsichten hinsichtlich des Vergleichs von Selbsteinschätzung und tatsächlichem Können führen, da diese mit Sicherheit, wie der Dunning-Kruger-Effekt besagt, voneinander abweichen. Weiters können die oben bereits angeführten, zusätzlich erhobenen Faktoren auch hier in die Forschung miteinbezogen werden: sexuelle Orientierung, Einfluss von Alkohol und Medikamenten oder der Einfluss vom schwankenden Hormonhaushalt bei Frauen, die hormonell verhüten oder bereits die Menopause gehabt haben. Vor allem nach der Beobachtung, dass auch das Befinden einen Einfluss auf die Selbsteinschätzung haben kann, könnten alle diese Items interessante Ergebnisse für die Forschung zu Tage bringen.

5. Literatur

- Alicke, M. D. (1985). Global Self-Evaluation as Determined by the Desirability and Controllability of Trait Adjectives. In *Journal of Personality and Social Psychology*, 49 (6), 1621 – 1630.
- Alicke, M. D. & Govorun, O. (2005). The Better-Than-Average Effect. In M. D. Alicke, D. A. Dunning & J. I. Krueger (Eds.), *Studies in self and identity. The Self in Social Judgment*, 85 – 106. New York: Psychology Press.
- Bateman, J. A. (1948). Intra-Sexual Selection in *Drosophila*. In *Heredity*, 2/3, 349 – 368.
- Bishop, G. F., Tuchfarber, A. J. & Oldendick, R. W. (1986). Opinions on fictitious issues: The pressure to answer survey questions. In *Public Opinion Quarterly*, 50, 240 – 250.
- Buunk, B.P. & Van Yperen, N. W. (1989). Social comparison, equality, and relationship satisfaction: Gender differences over a ten-year period. In *Social Justice Research*, 3, 157 – 180.
- Buunk, B. P. & van der Eijnden, R. J. J. M. (1997). Perceived Prevalence, Perceived Superiority, and Relationship Satisfaction: Most Relationships Are Good, but Ours is the Best. In *Personality and Social Psychology Bulletin*, 23 (3), 219 – 228.
- Caputo, D. D. & Dunning, D. (2005). What you don't know: The role played by errors of omission in imperfect self-assessments. In *Journal of Experimental Social Psychology*, 41, 488 – 505.
- Carter, T. J. & Dunning, D. (2008). Faulty self-assessment: Why evaluating one's own competence is an intrinsically difficult task. In *Personality and Social Psychology Compass*, 2, 346 – 360.
- Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?*, 73 – 96. Hillsdale: Erlbaum.
- Codol, J.-P. (1975). On the so-called “superior conformity of the self” behavior: Twenty experimental investigations. In *European Journal of Social Psychology*, 5 (4), 457 – 501.

- Dunning, D., Johnson, K., Ehrlinger, J. & Kruger, J. (2003). Why people fail to recognize their own competence. In *Current Directions in Psychological Science*, 12, 83 – 87.
- Dunning, D. (2011). The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology*, 44, 247 – 296.
- Eiser, J. R., Pahl, S. & Prins, Y. R. A. (2001). Optimism, Pessimism, and the Direction of Self–Other Comparisons. In *Journal of Experimental Social Psychology*, 37 (1), 77 – 84.
- Förstl, H. (2012). Theory of Mind: Anfänge und Ausläufer. In H. Förstl (Hrsg.) *Theory of Mind. Neurobiologie und Psychologie sozialen Verhaltens*, 2, 3 – 11. Berlin, Heidelberg: Springer.
- Galperin, A. & Haselton, M. G. (2013). Error management and the evolution of cognitive bias. In J. P. Forgas, K. Fiedler & C. Sedikides (Eds.), *Sydney symposium of social psychology. Social thinking and interpersonal behavior*, 45 – 63. New York: Psychology Press.
- Geary, D. C., Vigil, J. & Byrd-Craven, J. (2004). Evolution of Human Mate Choice. In *The Journal of Sex Research*, 41 (1), 27 – 42.
- Giladi, E. G. & Klar, Y. (2002). When standards are wide of the mark: Nonselective superiority and inferiority biases in comparative judgments of objects and concepts. In *Journal of Experimental Psychology: General*, 131, 538 – 551.
- Green, D. M. & Swets, J. A. (1966). Signal detection and psychophysics. New York: Wiley.
- Haselton, M. G. & Buss, D. M. (2000). Error Management Theory: A New Perspective on Biases in Cross-sex Mind Reading. In *Journal of Personality and Social Psychology*, 78, 81 – 91.
- Kaiser, M. K., Jonides, J. & Alexander, J. (1986). Intuitive reasoning about abstract and familiar physics problems. In *Memory and Cognition*, 14, 308 – 312.

- Klar, Y. & Giladi, E. E. (1997). No One in My Group Can Be Below the Group's Average: A Robust Positivity Bias in Favor of Anonymous Peers. In *Journal of Personality and Social Psychology*, 73 (5), 885 – 901.
- Kruger, J. (1999). Lake Wobegon Be Gone! The “Below-Average Effect” and the Egocentric Nature of Comparative Ability Judgments. In *Journal of Personality and Social Psychology*, 77, 221 – 232.
- Kruger, J. & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. In *Journal of Personality and Social Psychology*, 77 (6), 1121 – 1134.
- Kubesch, S. (2002). Sportunterricht: Training für Körper und Geist. In *Nervenheilkunde*, 9 (21), 487 – 490.
- Metcalf, J. (1998). Cognitive optimism: Self-deception or memory-based processing heuristics? In *Personality and Social Psychology Review*, 2 (2), 100 – 110.
- Nesse, R. M. (2006). The Smoke Detector Principle. In *Annals of the New York Academy of Sciences*, 935 (1), 75 – 85.
- Opdenacker, J., Delecluse, C. & Boen, F. (2009). The Longitudinal Effects of a Lifestyle Physical Activity Intervention and a Structured Exercise Intervention on Physical Self-Perceptions and Self-Esteem in Older Adults. In *Journal of Sport and Exercise Psychology*, 31 (6), 743 – 760.
- Otten, W. & Van der Pligt, J. (1966). Context effects in the measurement of comparative optimism in probability judgments. In *Journal of Social and Clinical Psychology*, 15, 80 – 101.
- Paulhus, D. L., Harms, P. D., Bruce, M. N. & Lysy, D. C. (2003). The over-claiming technique: Measuring self-enhancement independent of ability. In *Journal of Personality and Social Psychology*, 84, 890 – 904.
- Perloff, L. S., Fetzer, B. K. (1986). Self-other judgments and perceived vulnerability to victimization. In *Journal of Personality and Social Psychology*, 50, 502 – 510.

- Ross, L., Greene, D. & House, P. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. In *Journal of Experimental Social Psychology*, 13 (3), 279 – 301.
- Schkade, D. A. & Kahneman, D. (1998). Does living in California make people happy? A focusing illusion in judgments of life satisfaction. In *Psychological Science*, 9, 340 – 346.
- Sharot, T. (2011). The optimism bias. In *Current Biology*, 21 (23), R 941 – R945.
- Sonstroem, R. J. & Morgan, W. P. (1989). Exercise and self-esteem: Rationale and model. In *Medicine & Science in Sports & Exercise*, 21 (3), 329 – 337.
- Trivers, R. L. (1972). Parental Investment and Sexual Selection. In Campbell, B., *Sexual Selection and the Descent of Man*, 136 – 179. Chicago: Aldine Publishing Company.
- Van Yperen, N. W. & Buunk, B. P. (1990). A longitudinal study of equity and satisfaction in intimate relationships. In *European Journal of Social Psychology*, 20, 287 – 309.
- Van Yperen, N. W. & Buunk, B. P. (1991). Illusoire superioriteit het verband met belang en verifieerbaarheid van vergelijkingsdimensies [Illusory superiority: The relation with importance and verifiability of comparison dimensions]. In Van der Pligt, J., Van der Kloot, W., Van Knippenberg, A. & Poppe, M. A. (Eds.), *Fundamentele sociale psychologie*, 5, 186 – 200. Tilburg: Tilburg University Press.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. In *Journal of Personality and Social Psychology*, 39, 806 – 820.

Internetquellen

SurveyMonkey: <https://de.surveymonkey.com> [letzter Aufruf: 14.11.2019]

6. Anhang

Fragebogen

Nummer:	_____
Datum:	_____
Ort:	_____
Zeit:	_____

1. Alter: _____

2. Wie gut ist Ihr derzeitiges Befinden?

sehr schlecht ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ sehr gut

3. Haben Sie Erfahrung mit Geduldspielen?

Keine ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Viel

4. Wie gut sind Sie Ihrer Einschätzung nach in Geduldspielen?

sehr schlecht ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ sehr gut

5. Geschlecht: ☐ Weiblich ☐ Männlich ☐ Anderes:

6. *Wenn weiblich:*

➤ 1. Hatten Sie bereits ihre Menopause? ☐ Ja ☐ Nein

➤ 2. Wenn Ja: Nehmen Sie Hormonersatzprodukte ein?

Welche? _____

○ 3. Wenn Nein:

➤ Wie lange dauert Ihr durchschnittlicher Menstruationszyklus?

25 26 27 28 29 30 31 32 33 34 35 Tage

➤ Wann hat Ihre letzte Menstruation begonnen (Datum)? _____

➤ Nehmen Sie hormonelle Verhütungsmittel?

☐ Ja: _____

☐ Nein

7. Sexuelle Orientierung: ☐ Heterosexuell ☐ Homosexuell ☐ Bisexuell ☐ Anderes

8. Haben Sie derzeit einen Lebenspartner? ☐ Ja ☐ Nein

9. Stehen Sie unter dem Einfluss von Medikamenten?

☐ Ja: _____

☐ Nein

10. Haben Sie in den letzten 5 Stunden Alkohol konsumiert?

☐ Wenn Ja, wieviel? _____

☐ Nein

11. Haben Sie dieses Problem schon einmal gesehen? ☐ Ja ☐ Nein

12. Glauben Sie, dass dieses Problem grundsätzlich lösbar ist? ☐ Ja ☐ Nein

13. Glauben Sie, dass Sie dieses Problem in 2 Minuten lösen können? ☐ Ja ☐ Nein

_____ Spiel 1 _____

14. Glauben Sie, dass Sie dieses Problem mit mehr Zeit lösen könnten? ☐ Ja ☐ Nein

15. Glauben Sie, dass dieses Problem grundsätzlich lösbar ist? ☐ Ja ☐ Nein

16. Haben Sie dieses Problem schon einmal gesehen? ☐ Ja ☐ Nein

17. Glauben Sie, dass dieses Problem grundsätzlich lösbar ist? ☐ Ja ☐ Nein

18. Glauben Sie, dass Sie dieses Problem in 2 Minuten lösen können? ☐ Ja ☐ Nein

Spiel 2

19. Glauben Sie, dass Sie dieses Problem mit mehr Zeit lösen könnten? ☐ Ja ☐ Nein

20. Glauben Sie, dass dieses Problem grundsätzlich lösbar ist? ☐ Ja ☐ Nein

Anm. zur Spielabfolge: _____