

Die 100 Worte Textanalyse zur Messung impliziter Motive – eine Validierungsstudie

Autoren: Daniel Spitzer M.Sc. Psych., Leiter Forschung und Entwicklung 100 Worte; Dr. rer. nat. Stefan Dörr, geschäftsführender Partner von A47 Consulting; Helen op 't Roodt, freie Mitarbeiterin von A47 Consulting

Zusammenfassung

Ziel vorliegender Studie war die Überprüfung der Validität der 100 Worte Textanalyse durch Berechnung der Interrater-Reliabilitäten zwischen händischen und maschinellen Kodierungen von 1.100 Bildgeschichten der Picture Story Exercise (PSE). Es werden die Entwicklung eines automatisierten Verfahrens zur maschinellen Kodierung des Macht-, Beziehungs- und Leistungsmotivs und die Erstellung der 100 Worte Motiv-Korpora erläutert sowie bekannte Fehlerquellen in der Entwicklung einer regelbasierten Textanalyse und der Umgang mit diesen beschrieben. Ergebnisse zeigen mittlere Übereinstimmungen zwischen den händischen und maschinellen Kodierungen für das implizite Beziehungsmotiv ($n_{Aff} = 0.67$), das implizite Leistungsmotiv ($n_{Ach} = 0.6$) und etwas geringere Interrater-Reliabilitäten für das implizite Machtmotiv ($n_{Pow} = 0.47$). Die 100 Worte Textanalyse kann die drei Motivklassen valide messen und künftig auch der Untersuchung der impliziten Motivausprägung im Rahmen der beruflichen Entwicklung von Führungs- und Fachkräften dienen.

Einleitung

Motive können als unsere „Antreiber“ verstanden werden, die unser Erleben und Verhalten lenken. Dabei können explizite Motive, welche kognitiv verankert und bewusst zugänglich sind, von impliziten Motiven unterschieden werden, welche unbewusst unser Handeln und unsere Gefühle ausrichten und leiten (Hofer & Busch, 2011). Die impliziten Motive motivieren zum Ausüben einer Tätigkeit, weil die Ausführung selbst als befriedigend empfunden wird (Gröpel, Schoene & Wegner, 2015). Da implizite Motive nicht bewusst zugänglich und nicht verbal explizierbar sind (Brunstein, Dargel, Glaser, Schmitt & Spörer, 2008) und deswegen im Gegensatz zu den expliziten Motiven nicht durch Fragebögen oder Selbsteinschätzung erfasst werden können, werden sogenannte projektive Verfahren benötigt, um implizite Motive zu messen (Slabbinck, De Houwer, & Van Kenhove, 2011). In der Motivationsforschung war der Psychologe Henry Alexander Murray einer der Pioniere, der mit dem Thematischen Apperzeptions-Test (TAT) implizite Motive messbar machte. Als dahinter liegendes Prinzip wird angenommen, dass die Vorstellungen und Imaginationen in selbstverfassten Geschichten innere Bedürfnisse und implizite Handlungsmotive zum Ausdruck bringen. Diese Geschichten werden dann durch trainierte Kodierer nach einem validierten Kodiersystem (vgl. Smith, Feld & Franz 1992; Winter, 1994) inhaltsanalytisch ausgewertet.

Eine große Aufmerksamkeit wird der Messung drei zentraler Motive geschenkt, den sogenannten „Big Three“. Unterschieden werden hierbei das implizite Machtmotiv (*nPow*, z.B. das Bedürfnis, Einfluss auf andere Menschen auszuüben), das implizite Affiliations-/Beziehungsmotiv (*nAff*, z.B. das Bedürfnis, positive emotionale Beziehungen zu anderen aufzubauen und aufrecht zu erhalten) und das implizite Leistungsmotiv (*nAch*, z.B. das Bedürfnis, außergewöhnliche Leistungen zu erbringen) (Schultheiss & Brunstein, 2001). Tabelle 1 enthält eine nähere Beschreibung der drei Motivkategorien.

Tabelle 1: Beschreibung der drei zentralen impliziten Motive

Kontaktmotiv	Freude am geselligen Umgang mit anderen Menschen und Wunsch nach harmonischen Beziehungen; intuitives und spontanes Verhalten, darauf ausgerichtet vertrauensvolle Beziehungen aufzubauen, andere einzubinden und aktiv eigene Hilfestellung anzubieten.
Leistungs- motiv	Hohe emotionale Befriedigung, selbst besondere Leistungen zu vollbringen und mit hoher Ausdauer viel Arbeit und große Mühen auf sich zu nehmen, um anspruchsvolle Ziele und einen „Standard der Excellence“ zu erreichen.
Machtmotiv	Hohe Befriedigung „etwas zu bewegen und zu gestalten“; Bedürfnis, Verantwortung (u.a. auch Führungsaufgaben) zu übernehmen, andere zu beeinflussen und zu kontrollieren; keine Konflikte scheuen und sich mit Begeisterung für Visionen und höhere Werte einsetzen.

Besonders im Kontext von Führungsverhalten stellt das ganz persönliche Motivprofil einen wichtigen Erklärungswert dar (Furtner, 2012). So wurde das Leadership Motive Pattern (LMP) als eine besonders vielversprechende Konstellation der Motivausprägung identifiziert, welche sich aus einem hohen Bedürfnis nach Macht, geringer Ausprägung des impliziten Beziehungsmotivs und gleichzeitig hoher Activity Inhibition (Vorkommen des Wortes „nicht“) zusammensetzt (McClelland & Boyatzis, 1982). In neueren Studien wurde jedoch im Gegensatz zum klassischen LMP die Bedeutung eines höheren Affiliationsmotivs entdeckt, das die sozialen Einflussstrategien (*nPow*) durch eine offene, vertrauensvolle und partnerschaftliche Führungsbeziehung (*nAff*) erweitert und als Compassionate Leader Profile (CLP) bezeichnet wird (-vgl. Steinmann, Doerr, Schultheiss & Maier, 2015; -Steinmann, Ötting, & Maier, 2016). Spangler, Tikhomirov, Sotak und Palrecha (2014) untersuchten den Zusammenhang zwischen unterschiedlichen Mustern von Motivprofilen und dem Managementenerfolg in acht verschiedenen Formen von Organisationen (Miner, 1978; 1993).

Die Erhebung der Motivausprägung hat somit eine hohe Relevanz für die Erklärung, warum Menschen in verschiedenen Situationen auf eine bestimmte Weise handeln. Aufgrund des hohen ökonomischen Aufwands der händischen Kodierung von impliziten Motiven über die Picture Story Exercise (PSE) ist es ein großes Anliegen, die Kodierung von impliziten Motiven möglichst zu automatisieren (Blankenship, 2010), was erhebliche Vorteile im zeitlichen Aufwand und nicht zuletzt in der Transparenz der Bewertung bringt, da alle Bildgeschichten auf die gleichen Motivwörter hin untersucht und bewertet werden.

Vorliegende Studie hat das Ziel, die konvergente Validität der 100 Worte Motivkategorien mit der Picture-Story Exercise anhand von insgesamt 1.100 Bildergeschichten, welche von 186 Führungskräften verfasst worden sind, zu untersuchen. Die Studie soll die Ergebnisse der von Spitzer (2018) durchgeführten Validierungsstudie von 100 Worte replizieren. Ziel ist es, die Leistungsfähigkeit der automatisierten 100 Worte Motiv-Messung – und damit das Potenzial von automatischen Auswertungsverfahren allgemein - in verschiedenen Datensets aufzuzeigen.

Methode

Führungskräfte verschiedener deutscher Firmen und Non-Profit-Organisationen schrieben Bildgeschichten zu dem TAT-Bildmaterial (Dörr, 2008; Steinmann, Dörr, Schultheiss, & Maier, 2015). Als Anreiz erhielten die Führungskräfte differenziertes Feedback zu ihrer expliziten als auch impliziten Motivausprägung. Insgesamt wurden von 2005 bis 2019 von 205 Personen Geschichten ($N = 1.230$) zu den jeweils sechs unterschiedlichen Bildern gesammelt und 100 Worte zur Verfügung gestellt.

Zur Berechnung der konvergenten Validität zwischen der händischen und der maschinellen Kodierung über die 100 Worte Textanalyse wurden Bildgeschichten, die auf Englisch verfasst worden sind, deren Wortanzahl unter 20 lag, oder deren Motivklassifikation unrealistisch hohe Ausprägungen aufwiesen (Hinweis auf fehlerhafte Klassifikationen) aus dem zu analysierenden Datensatz entfernt ($N = 120$). Bei der Analyse wurden die Daten von 186 Personen ($\sigma = 134$, $\varphi = 53$) im Alter zwischen 22 und 64 Jahren ($M = 41$ Jahre; $SD = 8,5$ Jahre), mit insgesamt $N = 1.110$ Bildgeschichten berücksichtigt. Neben den Bildgeschichten wurden einige demographische Variablen erhoben, wie z.B. Alter und Geschlecht, Angaben zum Unternehmen und Charakteristiken der Führungsaufgabe der Individuen, welche für die Validierung der Maschine jedoch nicht weiter berücksichtigt wurden.

Bildmaterial und Bildgeschichten

Um die impliziten Motivausprägungen zu den "Big Three" (Bedürfnis nach Leistung, Macht und Beziehung) (McClelland, Koestner & Weinberger, 1989) zu ermitteln, erhielten die Führungskräfte zur Motivmessung sechs Bilder, welche bereits in zahlreichen früheren Forschungen verwendet worden sind (Smith, 1992; McClelland, 1975; McClelland & Steele, 1972). Präsentiert wurden die Bilder „Boxer“, „Women in laboratory“, „Ship captain“, „Couple by river“, „Trapeze artists“ und „Nightclub scene“. Bei 66 Führungskräften wurde anstatt des Bildes „Boxer“ das Bild „Architect at desk“ verwendet (Schultheiss & Brunstein, 2001). Die Bilder wurden intensiv in früheren Studien untersucht und so ausgewählt, dass sie populationsübergreifend motivationale Inhalte hervorrufen und gleichzeitig die Motivkategorien unterschiedlich, aber ausgewogen ansprechen (Pang, 2010).

Die Teilnehmer erhielten, wie von Schultheiss und Pang (2007) empfohlen, eine einheitliche Instruktion nach Winter (1992) und Lundy (1988). Während der Aufgabe wurde jedes Bild für jeweils

30 Sekunden dargeboten. Im Nachfolgenden hatten die Führungskräfte fünf Minuten Zeit, zu jeder auf dem Bild dargestellten Situation eine freie Geschichte zu schreiben und darüber, was als nächstes passieren könnte (Veroff, 1992; empfohlen von Smith, 1992). Die Bildgeschichten wurden teils händisch geschrieben ($N = 396$) und teils im Computer eingetippt ($N = 834$). Die handschriftlich verfassten Texte wurden nachträglich für die Datenauswertung mit der 100 Worte-Maschine transkribiert. Durchschnittlich enthielten die Texte, ähnlich wie in vorangegangenen Studien bereits beschrieben (Pang, 2010), etwa 100 Worte ($M = 103$; $SD = 53,9$) mit einer Spanne von 22 bis zu 607 Wörtern.

Kodierregeln

David Winter integrierte 1994 das *Manual for Scoring Motive Imagery in Running Text*, welches Regeln und Instruktionen zur motivbezogenen inhaltsanalytischen Auswertung von geschriebenen oder gesprochenen Fließtexten nahezu jeder Art (bspw. Reden, Literaturarbeiten, oder Interviews) enthält. In diesem Manual definiert Winter „motive image“ als eine Handlung, ein Wunsch, Anliegen, oder jegliche andere Form von innerem Zustand, welchen eine Person sich selbst oder einer, oder mehreren anderen Personen zuschreibt. Das Auswertungssystem umfasst das Leistungs- und das Machtmotiv sowie eine kombinierte Kategorie des Affiliations- und des Intimitätsmotivs, welche in der empirischen Überlappung beider Motivformen begründet ist. Die wichtigsten und für im Nachfolgenden relevantesten Regeln für die Kodierung für PSE sind hier zusammengefasst (S. 22f.):

1. Ein Motiv kann nur einmal in einem einzelnen Satz gewertet werden, in einem Satz können jedoch mehrere verschiedene Motive berücksichtigt werden (Ausnahme s. Regel 3).
2. Tritt ein Motiv in zwei separaten, aber nacheinander folgenden Sätzen auf, wird es nur einmal gewertet; wenn ein Motiv im ersten und dritten Satz vorkommt, wird es im ersten und im dritten Satz kodiert.
3. Wenn ein Motiv mehrmals in einem Satz oder in den nachfolgenden Sätzen vorkommt, aber durch ein weiteres Motiv getrennt wird, wird es mehrmals kodiert (z.B. Leistung – Macht – Leistung).

Motivbilder, die verneint werden, werden nicht kodiert (z. B. „Sie hat ihn nicht getötet“ wird *nicht* als Machtmotiv kodiert). Des Weiteren werden Motive bei einfachen Fragen nicht kodiert, während sie bei hypothetischen, rhetorischen oder instrumentellen Fragen hingegen doch kodiert werden. Generell sollten Kodierungen auf Wörtern und Phrasen und nicht auf deren intuitiven Interpretation basieren. Das Heranziehen des Kontextes um ein Wort oder einen Satz zu verstehen ist jedoch erlaubt.

Händische Kodierung

Die Geschichten wurden für $nPow$, $nAch$ und $nAff$ jeweils von zwei gut ausgebildeten Kodierern kodiert, welche eine Übereinstimmungsgüte von über .85 mit dem Kalibrierungsmaterial in dem

Kodierungshandbuch von Winter (1994) für implizite Motive aufwies. Kodierer unterschieden sich zwischen den Stichproben, sodass an dem gesamten Datensatz fünf trainierte Kodierer mitgewirkt haben. Die Interrater-Reliabilität zwischen den Kodierern bei einer Textauswahl von $N = 207$ zeigt gute bis sehr gute Übereinstimmungen auf (berechnet für drei Stichprobengruppen: $r_1 = 0.884$; $r_2 = 0.952$; $r_3 = 0.929$ wobei Anzahl der Texte $N_1 = 75$, $N_2 = 90$, $N_3 = 42$ beträgt). Diskrepanzen in der Kodierung wurden zwischen den beteiligten Kodierern kommunikativ validiert und es wurde ein Konsens gefunden. Durchschnittlich wurden für $nPow$, $nAff$ und $nAch$ jeweils ein Motiv für Leistung, Affiliation und Macht pro Bild gefunden. Des Weiteren wurde bei der händischen Kodierung die Ausprägung der sog. „Activity Inhibition“ gemessen, in dem alle Verneinungen durch das Wort „nicht“ (McClelland, 1979) gezählt worden sind.

Kodierung mit der 100 Worte Textanalyse – Die 100 Worte Motiv-Korpora

Die 100 Worte Textanalyse enthält Wörterbücher zur Erfassung impliziter Motive. Diese teilen sich gemäß der Definition von McClelland (1987)⁷ in Wörterbücher auf, die das Machtmotiv, das Leistungsmotiv sowie das Beziehungsmotiv erfassen. Jede dieser drei Motivkategorien unterteilt sich in drei Unterkategorien (Annäherungstendenz, Vermeidungstendenz, sowie eine neutrale, valenzfreie Kategorie). Alle Motiv-Wörterbücher umfassen 7438 Worte.

Zur Veranschaulichung soll im Folgenden auf die Konstruktionsweise der 100 Worte Motive-Korpora eingegangen werden. Das Vorgehen bei der Erstellung lässt sich in vier Phasen gliedern, die sich aus manuellen und Data Science Methoden zusammensetzen. In der anfänglichen Initialisierungsphase wurden durch Brainstorming zu den Motivkategorien semantisch-assoziierte Begriffe gesammelt. Diese Assoziationen dienten als Grundlage für alle weiteren Schritte. In der zweiten Phase wurden mithilfe der Kosinus-Ähnlichkeit ähnliche Worte zu den bereits bestehenden hinzugefügt. Im dritten Schritt wurde die Basisraten-Häufigkeit aller gefundenen Wörter erhoben und diese nach Häufigkeit sortiert. Die Wörter mit höchstem Vorkommen standen ganz oben, die mit dem niedrigsten ganz unten. Schließlich wurde die interne Konsistenz der erstellten Wörterbücher berechnet und durch schrittweise Wegnahme einzelner Wörter untersucht, welchen Beitrag sie zur gesamten Konsistenz leisten. Bei der Entscheidung über die Auswahl der Begriffe für das finale Wörterbuch wurden Schwellenwertkriterien der oben beschriebenen statistischen Kenngrößen wie die Ähnlichkeit, die Basisrate und der Beitrag zur Konsistenz genutzt. Die so entstandenen Wörterlisten wurden von drei Personen auf Passung der Begriffe zur jeweiligen Motivkategorie überprüft. Um in das Wörterbuch aufgenommen zu werden, wurde die Zustimmung von mindestens zwei Annotatoren benötigt. Bei nicht eindeutig zuordbaren Begriffen handelte es sich häufig um ambige Wörter.

Ambige Worte sind solche, die mehrere Bedeutungen tragen können. So trägt beispielsweise das Wort „nötigen“ im Kontext „Ich benötige dich“ eine Machtkomponente, im Kontext „schicken Sie

bitte die nötigen Unterlagen“ hingegen nicht. Wenn „nötigen“ in das Wörterbuch „Macht und Führung“ aufgenommen werden würde, dann würde die Textanalyse fälschlicherweise das Wort in allen Kontexten als Machtwort klassifizieren. Schwartz und Kollegen, (2013) identifizierten als Fehlerquellen bei regelbasierten Textanalysen die Fehlklassifizierung komplexer Negationen, Ironie, falsche Autorenschaft, Rechtschreibfehler und am häufigsten, die Verwendung falscher Bedeutungen. Bei der Erstellung der Wörterbücher wurde deshalb darauf geachtet, dass Fehler erster Art (durch Nicht–Aufnahme eines Wortes, um keine False-positives zu produzieren), sowie zweiter Art (durch Aufnahme, um keine False-negatives zu produzieren) bestmöglich reduziert werden. Dazu wurde ein maschinelles Verfahren genutzt, das den Kontext eines Wortes berücksichtigt.

Mit einer Mischung aus maschinellem Lernen und regelbasierten Verfahren wurden weitere Fehlerquellen identifiziert, wie zu dem Umgang mit Wortstambildung, Verstärker-Präposition, Negationen und Ngrammen. Darauf möchten wir im Folgenden eingehen.

Wortstamm-Bildung/ Stemming	Einige regelbasierte Textanalysen (z. B. LIWC) kürzen aus zeitökonomischen- und Effizienzgründen verschiedene Formen von Wörtern (z.B. Deklinationen des Verbs „gehen“) auf einen gemeinsamen Wortstamm und nehmen nur diesen in das Wörterbuch auf. Schwartz und Kollegen (2013) fanden jedoch, dass dadurch die Rate von Fehlklassifikationen ansteigt. Durch Stemming mit dem NLTK-Stemmer wird (http://www.nltk.org/howto/stem.html) beispielsweise das Wort „Note“ auf „Not“ gekürzt und so zu einem ambigen Wort. Aus diesem Grund wird für unsere Analyse kein Stemming verwendet.
Verstärker-Präposition	Adjektive mit einer verstärkenden Präposition (z. B. „sehr schön“) werden von der 100 Worte Textanalyse erfasst und bekommen ein stärkeres Gewicht (1,2-fache Wertung) zugeschrieben.
Negationen	Einige Kategorien in der 100 Worte Textanalyse drücken eine Wertung und somit eine Polarität aus. Da Negationen die Bedeutung eines Wortes umkehren können, wurde ein Verfahren entwickelt, welches die häufigsten Formen von Verneinungen entdeckt (d.h. Negation erscheint direkt oder unmittelbar vor der in Frage stehenden Valenz, z.B.: „nicht schön“ bzw. „nicht sehr schön“). Bei den Motivkategorien werden negierte Wörter im Text, die einer Motiv-Valenz-Kategorie entstammen (z. B. Positive Achievement), der jeweils gegenteiligen Kategorie zugerechnet (z. B. Negative Achievement).
Ngramme	Da sich Bedeutungen manchmal nur in der Kombination von Worten ausdrücken, werden sogenannte Ngramme, also Wortkombinationen, die aus zwei oder mehr Worten bestehen, gesammelt und in den Wörterbüchern aufgenommen. Im Satz „Mit dir gehe ich durch dick und dünn.“ drückt das 4-Gramm „durch dick und dünn“ eine Bedeutung (z. B. Beziehung) aus, die die einzelnen Substituenten alleine oder in einer anderen Kombination nicht hätten.

Datenaufbereitung für den Vergleich von händischer und maschineller Kodierung

Zur Analyse wurden die 186 PSE in die sechs zugrundeliegenden Geschichten (N = 1.100) aufgeteilt, um die einzelnen Geschichten – und nicht die gesamte Übung – zu analysieren. Um die Motivklassifikation mit der Analyse von 100 Worte besser vergleichen zu können, wurden die absoluten Häufigkeiten der gefundenen Motive an der Wortanzahl der Texte relativiert und mit der relativen Häufigkeit weitergearbeitet. Bei der Datenaufbereitung war auffallend, dass die Mittelwerte der händisch kodierten Motive in verschiedenen Abschnitten des Datensatzes signifikant

unterschiedlich waren. Das bedeutet also, dass die Neigung, ein bestimmtes Motiv zu kodieren, nicht überall gleich hoch war. Dieser Unterschied konnte auf die Kodierkräfte zurückgeführt werden. Obwohl die Zusammenhänge von etwa $r = 0.74$ bis $r = 0.86$ recht hoch sind, weichen sie trotzdem vom Ideal einer vollkommenen Interrater-Reliabilität ab. Gerade diese Schwäche der manuellen Kodierung macht die Notwendigkeit einer Alternative, wie etwa automatische Verfahren, deutlich, die aufgrund ihrer Konstruktionsweise kein Problem mit der Interrater-Reliabilität haben. Aus diesem Grund wurde der Datensatz in Abhängigkeit des Kodierers geclustert und im Weiteren mit den sieben so entstandenen Clustern gearbeitet. Hätten wir mit dem nicht-geclusterten Datensatz weitergearbeitet, wäre die Linearität beider Merkmale (abhängige und unabhängige Variable)-, die Bedingung für die Berechnung der Pearson-Korrelation ist, wahrscheinlich nicht erfüllt gewesen.

Ergebnisse

Die Ergebnisse der deskriptiven Statistik sind Tabelle 1 zu entnehmen. Für den Kolmogorow-Smirnow-Test zeigt sich eine Abweichung der Motiv-Kategorien von der zu erwartenden Normalverteilung. Bei einem Signifikanzniveau von $\alpha = .05$ und $N = 1.110$ Beobachtungen darf die größte absolute Abweichung den Wert von $\alpha = .041$ nicht überschreiten. Alle drei Motivkategorien überschreiten diesen Wert (Leistungsmotiv: .167; Machtmotiv: .182; Beziehungsmotiv: .131). Auch Schultheiss (2013) beobachtete bei seiner Analyse der LIWC Wörterbücher ein Abweichen von der Normalverteilung und empfahl daher eine log-Transformation ($\text{newScore} = \ln(1 + \text{oldScore})$). Wir folgten diesem Vorschlag und führten die Transformation mit allen Kategorien durch.

Tabelle 1. Deskriptive Statistik (M=Mittelwert; SD=Standardabweichung) und interne Konsistenz (=Cronbach's Alpha) für die manuell kodierten Motive (overall) und die 100 Worte Motivkategorien.

	<i>Alpha</i>	<i>M</i>	<i>SD</i>	<i>Range</i>
<i>PSE [Dörr Datensatz]</i>				
<i>n Macht</i>	0.36	1.31	1.31	0 -12
<i>n Leistung</i>	0.51	1.00	1.26	0 - 10
<i>n Beziehung</i>	0.48	1.00	1.34	0 - 7
<i>100 Worte</i>				
<i>Macht</i>	0.44	2.04	2.25	0 - 16
<i>Leistung</i>	0.54	2.89	3.00	0 - 17
<i>Beziehung</i>	0.49	3.85	3.44	0 - 19

Anmerkung. Die Angaben beziehen sich auf den gesamten (und nicht auf den geclusterten) Datensatz

Um Aufschluss über die konvergente Validität zu erhalten, wurden Korrelationen nach Pearson berechnet. Es zeigt sich ein signifikanter Zusammenhang zwischen den Ergebnissen der 100 Worte Motivkategorien und den manuell-kodierten Motiven. Alle drei Motivkategorien korrelieren in einem mittleren bis hohen Maß positiv mit den manuell-kodierten Motiven (s. Tabelle 3).

Tabelle 2. Korrelationen zwischen traditionell kodierten Motiven und den 100 Worte Motivkategorien

<i>100 Worte Kategorie</i>	<i>Transformation</i>	<i>n Beziehung</i>	<i>n Leistung</i>	<i>n Macht</i>
<i>n Beziehung</i>	In	0.67 (0.59)		
<i>n Leistung</i>	In		0.60 (0.54)	
<i>n Macht</i>	In			0.47 (0.38)

Anmerkung. Alle Ergebnisse sind auf dem 0.001 Niveau signifikant. In Klammern finden sich die Korrelationswerte für den gesamten Datensatz

Zu beachten ist, dass die Werte in Tabelle 2 die Mittelwerte aus den sieben Clustern darstellen. Die Korrelationswerte über den gesamten Datensatz sind in Klammern aufgeführt.

Diskussion

Zielsetzung dieser Studie war es, zu zeigen, dass die automatisierte Erfassung von Worten und Wortkombinationen mit der 100 Worte Textanalyse Motive vorgenommen werden kann. Dazu wurden die zur Verfügung gestellten PSE-Texte von Dr. Stefan Dörr, A47-Consulting, mit den 100 Worte Motivkategorien untersucht und anschließend die konvergente Validität berechnet.

Wie bereits erwähnt, sind thematische Inhaltsanalysen zeitintensiv und mit viel Arbeit verbunden. Ein erfahrener Kodierer braucht ca. 2 bis 5 Minuten um eine PSE Geschichte zu kodieren, sodass die Kodierung von Geschichtsmaterial für sechs Bilder pro Teilnehmer etwa 12 bis 30 Minuten in Anspruch nimmt (Pang, 2010). Die maschinelle Kodierung mit der 100 Worte Textanalyse untersucht einen Text in wenigen Sekunden auf Motivwörter und sorgt somit für eine erhebliche Zeitersparnis.

Als weiterer Kritikpunkt an der manuellen Kodierung der PSE ist die eingeschränkte Transparenz eines projektiven Verfahrens zu nennen. Häufig wird Teilnehmern an einem PSE kein Einblick in die Bewertung ihrer Bildgeschichten gegeben, die 100 Worte Analyse kennzeichnet die verschiedenen Motivwörter farbig (siehe Abbildung 1).

Außerdem ist eine maximale Gleichförmigkeit der Kodierung gegeben, da die Maschine, anders als Menschen, die in die Motiv-Korpora eingespeisten Wörter immer auf die gleiche Weise kodiert (es sei denn, die Motivwörter werden verneint) und nicht wie bei der manuellen Kodierung personenabhängig ist, also Unterschiede durch verschiedene Rater zustande kommen.

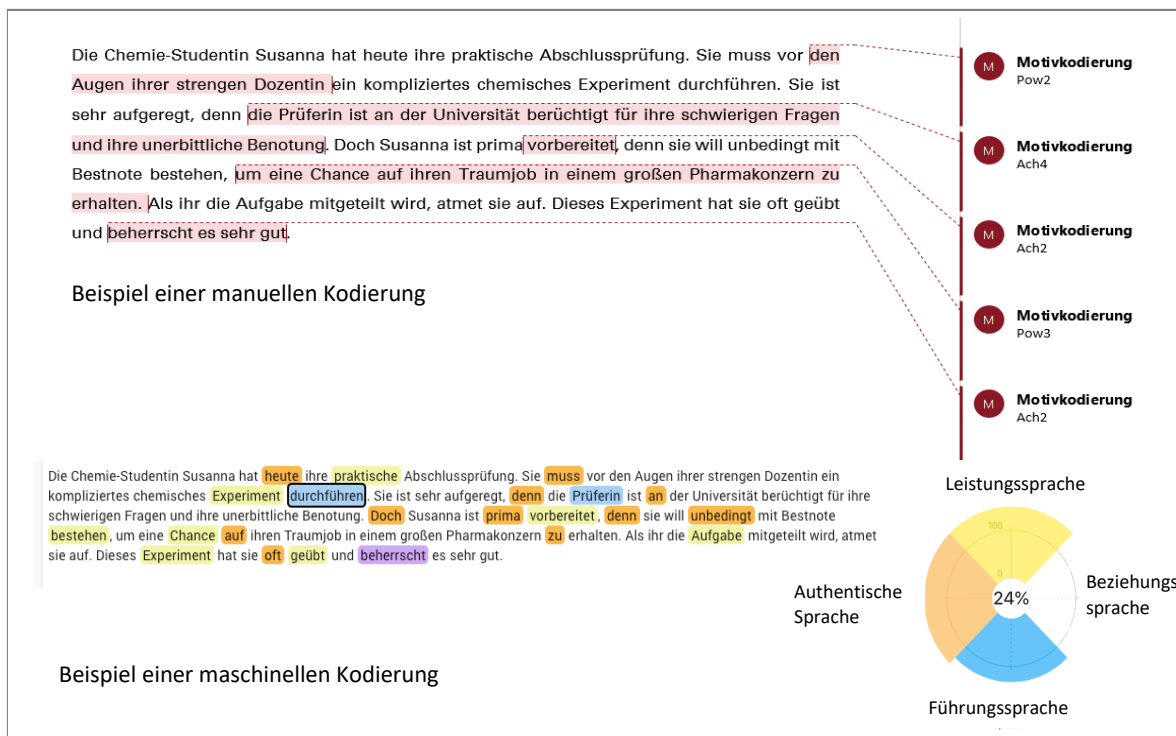


Abb. 1: Vergleich der Darstellungsweise der manuellen und maschinellen Kodierung

Limitationen

Angesichts der durchaus positiven Ergebnisse der vorliegenden Studie sollen aber auch einige Aspekte kritisch beleuchtet werden und zu weiteren Untersuchungen stimulieren. So kann die Maschine beispielsweise nicht zwischen rhetorischen (hier werden Motive kodiert) und einfachen Fragen (hier werden Motive nicht kodiert) unterscheiden. Da in den Bildgeschichten sowieso nur selten Fragen vorkommen, ist dieser Kritikpunkt aber im weitesten Sinne vernachlässigbar. Außerdem wird, im Gegensatz zu der nach Winter (1994) empfohlenen Kodierung, jedes einzelne Motivwort kodiert – ohne Rücksicht darauf, ob dasselbe Motiv in demselben Satz bereits kodiert wurde. Das erklärt auch die durchschnittlich höhere Anzahl an Motiven bei der maschinellen Kodierung. Dabei stellt sich jedoch die Frage, ob diese Methode der Kodierung nicht auch seine Berechtigung hat, denn man könnte auch argumentieren, dass ein häufigeres Vorkommen und sehr motivreiche Sätze nicht auch stärker in die Gewichtung mit einfließen sollten.

Niedrigere Interrater-Reliabilitäten wurden über den gesamten Datensatz hinweg für das Machtmotiv gefunden. Im Allgemeinen wird auch in der Forschung die inhaltliche Zusammensetzung des Machtmotives diskutiert und seine Definition unterscheidet sich zwischen unterschiedlichen Motivpsychologen (für eine nähere Erläuterung s. Veroff, 1957; Uleman, 1966; McClelland, 1970; Schultheiss & Brunstein, 2005). Allein aus historischer Sicht kann man bezüglich des Machtmotives somit Uneinigigkeiten entdecken, je nachdem, nach welchem Konzeptverständnis man vorgeht. Eine Möglichkeit, um die automatisierte Erfassung des impliziten Machtmotivs zu optimieren, wäre eine semantische Erweiterung des bestehenden Wörterbuchs zum Machtmotiv um zentrale Dimensionen

einflussorientierter Führungstheorien, wie z.B. Transformationale, Charismatische, oder Authentische Führung (vgl. Felfe, 2015).

Was sich bei unseren Analysen der PSE-Geschichten immer wieder zeigte, ist die Interpersonalität des Machtmotivs, gemäß der Definition von Winter (1994) in seinem Kodierungsmanual: „Strong, forceful actions that inherently have impact on others“ (S. 3). Diese Interpersonalität der Einflussnahme drückt sich auch sprachlich aus. So enthält der Satz „*Sie erklärt ihm, was er machen soll.*“ eindeutig ein Machtmotiv (*nPow*), weil zwei Personen in einem (Einfluss-)Verhältnis stehen. Wäre nur von einer Person die Rede, wäre es hier nicht angebracht ein Machtmotiv zu kodieren (z. B. „*Sie erklärt sich die Zusammenhänge im Modell*“). Da jedoch kein Macht-typisches Inhaltswort enthalten ist, ist es schwierig, die granularen Bedeutungsunterschiede des Wortes „erklären“ sowie die Abhängigkeiten zwischen Entitäten in Texten maschinell korrekt zu erfassen. Zwar gibt es Ansätze zur Erfassung von Abhängigkeiten (Dependency Parsing), doch funktionieren diese nicht mit der erwünschten Qualität, wie eigene Benchmark-Tests gezeigt haben. Die Entwicklung von Verfahren zur Erfassung von Entitäten, die in einem (Macht-)Verhältnis stehen, wäre daher ein wichtiger Fortschritt von automatisierten Motivtests.

Ausblick zur weiteren Anwendung in der Praxis

Ein weiteres Maß, welches in der Motivforschung im Rahmen der Picture Story Exercise viel Aufmerksamkeit erhält, ist die sogenannte „Activity Inhibition“. Darunter wird die stabile Tendenz verstanden, emotionale und motivationale Impulse zu hemmen (McClelland, Davis, Kallin & Wanner, 1972; Langens, 2010). Die 100 Worte Textanalyse konzeptualisiert Negationen von Motiven nicht als Activity Inhibition, sondern rechnet sie der gegenteiligen Kategorie zu. Vorteil dieses Verfahren ist, dass so die Inhibition eines negativen Motiv-Wortes erfassbar wird. So könnte beispielsweise aus der Formulierung „nicht schlecht“ ein positives Achievement-Motiv gedeutet und kodiert werden.

Weiterhin kann es als sinnvoll erachtet werden, auf Basis von Ergebnissen zur impliziten und expliziten Motivausprägung Übereinstimmungen und Diskrepanzen zwischen den selbst zugeschriebenen Motiven (explizit) und den unbewussten (impliziten) Motiven zu finden (McClelland et al., 1989). Je höher die Übereinstimmung ist, desto leichter können explizite und implizite Motive als psychische Ressourcen und Kraftquellen genutzt werden, das eigene Handeln wird dadurch effizienter und es steigt das subjektive Wohlbefinden (Wagner, Baumann & Hank, 2016). Je höher die Diskrepanz ist, desto mehr "innerer Kraftaufwand" ist für die Regulation und Ausbalancierung zwischen „Verstand und Gefühl“ notwendig (Kehr, 2004).

Die vorliegende Studie zeigt, dass implizite Motive durch eine automatisierte Kodierung durch die 100 Worte Textanalyse valide gemessen werden können. Dadurch sind wir in der Lage, für die Diagnostik und Entwicklung von impliziten Motiven (sowohl über das Testverfahren PSE -Picture Story

Exercise, als auch –über andere sprachliche Datenquellen, wie z.B. Interviews, Reden, Workshop-Dokumentationen etc.) eine zuverlässige, automatisierte und damit ökonomische Methodik der Motivmessung für die Praxis in Unternehmen anzubieten.

Im Rahmen der Personal- und Organisationsentwicklung von Unternehmen empfehlen wir deshalb für die berufliche Entwicklung von Führungs- und Fachkräften sowohl implizite und explizite Motive durch Motive-Potentialanalysen zu diagnostizieren, als auch gezielt im Rahmen von geeigneten Coaching- & Trainingsinterventionen mit motivpsychologisch qualifizierten Beratern zu entwickeln (vgl. Dörr, Schmidt-Huber, Inderst & Maier, 2015; Dörr, Hund, & Inderst, 2016). Je besser Führungskräfte und Mitarbeiter in ihrem beruflichen Kontext entsprechend der eigenen Motivdispositionen eingesetzt werden und handeln können, desto stärker engagieren sie sich auch für Ziele, lernen Neues, entwickeln eigene Kompetenzen weiter, arbeiten konstruktiv zusammen und können Krisen besser bewältigen - und "fühlen sich in der eigenen Arbeit motiviert".

Literaturverzeichnis

- Blankenship, V. (2010). Computer-based modeling, assessment, and coding of implicit motives. In O. Schultheiss & J.C. Brunstein (Hrsg.), *Implicit motives* (S. 186-208). New York: Oxford University Press.
- Brunstein, J. C., Dargel, A., Glaser, C., Schmitt, C. H., & Spörer, N. (2008). Persönliche Ziele im Studium: Erprobung einer Intervention zur Steigerung der Zieleffektivität und Zufriedenheit im Studium. *Zeitschrift für Pädagogische Psychologie*, 22(34), 177-191.
- Dörr, S. (2008). *Motive, Einflussstrategien und transformationale Führung als Faktoren effektiver Führung*. München: Rainer Hampp Verlag.
- Dörr, S. L., Schmidt-Huber, M., Inderst, F., & Maier, G. W. (2015). Führungskompetenzen diagnostizieren und entwickeln: Anwendung des LEaD-Kompetenzmodells effektiver Führung in der Praxis. In R. van Dick & J. Felfe (Eds.), *Handbuch Mitarbeiterführung: Wirtschaftspsychologisches Praxiswissen für Fach- und Führungskräfte*. Berlin: Springer.
- Dörr, S., Hund, A., & Inderst, F. (2016). Motiv-Profil-Analyse: Ein wirksames Instrument für die Diagnose und Entwicklung von Führungskompetenzen im Wissenschaftskontext. *Personal- und Organisationsentwicklung*, 2, 47-54. UVW Verlag.
- Felfe, J. (2015). *Trends der psychologischen Führungsforschung: Neue Konzepte, Methoden und Erkenntnisse*. Göttingen: Hogrefe Verlag.
- Furtner, M. R. (2012). Wie beeinflussen Motive das Führungsverhalten. *Psychologie des Alltagshandelns*, 5(2), 52-65.
- Gröpel, P., Schoene, L., & Wegner, M. (2015). Implizite und explizite Motive von Leistungs- und Freizeitsporttreibenden. *Zeitschrift für Sportpsychologie*, 22(1), 6-19.
- Hofer, J. & Busch, H. (2011). Satisfying one's needs for competence and relatedness: Consequent domain-specific well-being depends on strength of implicit motives. *Personality and Social Psychology Bulletin*, 37(9), 1147-1158.
- Kehr, H. M. (2004). Implicit/explicit motive discrepancies and volitional depletion among managers. *Personality and Social Psychology Bulletin*, 30(3), 315-327.
- Langens, T. A. (2010). Activity inhibition. In O. Schultheiss & J.C. Brunstein (Hrsg.), *Implicit motives* (S. 89-115), New York: Oxford University Press.
- Lundy, A. (1988). Instructional set and Thematic Apperception Test validity. *Journal of Personality Assessment*, 52, 309-320.
- McClelland, D. C. (1970). The two faces of power. *Journal of International Affairs*, 24, 29-47.
- McClelland, D. C. (1975). *Power: The inner experience*. New York: Irvington Publishers.
- McClelland, D. C. (1987). *Human Motivation*. New York: Cambridge University Press.

- McClelland, D. C. (1979). Inhibited power motivation and high blood pressure in men. *Journal of Abnormal Psychology, 88*, 182–190.
- McClelland, D. C., & Boyatzis, R. E. (1982). Leadership motive pattern and long-term success in management. *Journal of Applied psychology, 67*(6), 737.
- McClelland, D. C., Davis, W. N., Kalin, R., & Wanner, E. (1972). *The drinking man: Alcohol and human motivation*. New York: Free Press.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690–702.
- McClelland, D. C., & Steele, R. S. (1972). *Motivational workshops*. New York: General Learning Press.
- Miner, J.B. (1978). Twenty years of research on role-motivation theory of managerial effectiveness. *Personnel Psychology, 31*(4), 739–760.
- Miner, J.B. (1993). *Role motivation theories*. London: Routledge.
- Pang, J. S. (2010). Content coding methods in implicit motive assessment: Standards of measurement and best practices for the picture story exercise. *Implicit motives, 1*, 119-151.
- Schultheiss, O.C. & Brunstein, J.C. (2001). Assessment of Implicit Motives with a Research Version of the TAT: Picture Profiles, Gender Differences, and Relations to Other Personality Measures. *Journal of Personality Assessment, 77*(1), 71-86.
- Schultheiss, O.C. & Brunstein, J.C. (2005). An implicit motive perspective on competence. In A. J. Elliot & C. C. Dweck (Hrsgs.), *Handbook of competence and motivation* (pp. 31-51). New York: Guilford Publications.
- Schultheiss, O.C. & Pang, J.S. (2007). Measuring implicit motives. In R.W. Robins, R.C. Fraley & R. F. Krueger (Hrsg.), *Handbook of research methods in personality psychology* (pp. 322-344). New York: Guilford Press.
- Slabbinck, H., De Houwer, J., & Van Kenhove, P. (2011). A pictorial attitude IAT as a measure of implicit motives. *European Journal of Personality, 25*(1), 76-86.
- Smith, C. P. (1992). *Motivation and personality: Handbook of thematic content analysis*. Cambridge: Cambridge University Press.
- Spangler, W. D., Tikhomirov, A., Sotak, K. L., & Palrecha, R. (2014). Leader motive profiles in eight types of organizations. *The Leadership Quarterly, 25*(6), 1080-1094.
- Schultheiss, O.C. (2013). Are implicit motives revealed in mere words? Testing the marker-word hypothesis with computer-based text analysis. *Frontiers in Psychology, 4*, pp. 1-20.
- Schwartz, H. A., Eichstaedt, J., Blanco, E., Dziurzyński, L., Kern, M. L., Ramones, S., ... & Ungar, L. (2013, June). Choosing the right words: Characterizing and reducing error of the word count approach. In *Second Joint Conference on Lexical and Computational Semantics (*SEM)*, Volume

- 1: Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity (pp. 296-305).
- Spitzer, D. (2018). Lassen sich Ergebnisse aus impliziten Motivtests mit den 100 Worte Motivkategorien vorhersagen: Eine Validierung. [Weblog] https://www.100worte.de/wp-content/uploads/2019/07/100Worte_Validierung_Motive.pdf Zugriff am: 25.10.2019
- Steinmann, B., Dörr, S. L., Schultheiss, O. C., & Maier, G. W. (2015). Implicit motives and leadership performance revisited: What constitutes the leadership motive pattern?. *Motivation and Emotion*, 39(2), 167-174.
- Steinmann, B., Ötting S. K. & Maier G. W. (2016). Need for affiliation as a motivational add-on for leadership behaviors and managerial success. *Front. Psychol.*, 7, 1972.
- Uleman, J. S. (1966). A new TAT measure of the need for power. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.
- Veroff, J. (1957). Development and validation of a projective measure of power motivation. *The Journal of Abnormal and Social Psychology*, 54(1), 1-8.
- Veroff, J. (1992). Thematic apperceptive methods in survey research. In C. P. Smith (Ed.) *Motivation and personality: Handbook of thematic content analysis* (S. 100-109). New York: Cambridge University Press.
- Wagner, L., Baumann, N., & Hank, P. (2016). Enjoying influence on others: Congruently high implicit and explicit power motives are related to teachers' well-being. *Motivation and Emotion*, 40(1), 69-81.
- Winter, D. G. (1992). Power motivation revisited. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic content analysis* (pp. 301-310). New York: Cambridge University Press.
- Winter, D.G. (1994). *Manual for scoring motive imagery in running text*. Unpublished instrument, University of Michigan, Ann Arbor.