

FULL PAPER

Skala Spannungserleben (SSE). Entwicklung einer Skala zur postrezeptiven Erfassung von Spannungserleben während der Rezeption narrativer Inhalte

**Skala Spannungserleben (SSE) [scale for measuring suspense (SSE)].
A scale for the post exposure measurement of narrative suspense**

Carsten Wunsch, Theresa Briselat, Sophie Reitmeier, Jana Keil & Miriam Czichon

Carsten Wünsch (Prof. Dr.), Universität Bamberg, Institut für Kommunikationswissenschaft, An der Weberei 5, 96047 Bamberg, Germany. Kontakt: carsten.wuensch(at)uni-bamberg.de. ORCID: <https://orcid.org/0009-0002-5994-4463>

Theresa Briselat (M.A.), Universität Bamberg, Institut für Kommunikationswissenschaft, An der Weberei 5, 96047 Bamberg, Germany. Kontakt: thesa.briselat(at)uni-bamberg.de. ORCID: <https://orcid.org/0009-0000-8482-4530>

Sophie Reitmeier (M.A.), Universität Bamberg, Institut für Kommunikationswissenschaft, An der Weberei 5, 96047 Bamberg, Germany. Kontakt: sophie.reitmeier(at)uni-bamberg.de. ORCID: <https://orcid.org/0009-0005-7351-0890>

Jana Keil (M.A.), Universität Bamberg, Institut für Kommunikationswissenschaft, An der Weberei 5, 96047 Bamberg, Germany. Kontakt: jana.keil(at)uni-bamberg.de. ORCID: <https://orcid.org/0009-0004-6181-9084>

Dr. Miriam Czichon, Contact: miriam.czichon1(at)gmail.de.



Skala Spannungserleben (SSE). Entwicklung einer Skala zur postrezeptiven Erfassung von Spannungserleben während der Rezeption narrativer Inhalte

Skala Spannungserleben (SSE) [scale for measuring suspense (SSE)]. A scale for the post exposure measurement of narrative suspense

Carsten Wunsch, Theresa Briselat, Sophie Reitmeier, Jana Keil & Miriam Czichon

Zusammenfassung: „Spannung ist, wenn’s spannend ist“ (Vorderer, 1994). Über viele Jahre hinweg prägte dieser Satz sinngemäß die empirische Erforschung des – für das Unterhaltungserleben zentrale – Konstrukts *Spannungserleben*. Postrezeptiv erfasst, wurde Spannung in der Regel eindimensional oder als eine Subdimension anderer Konstrukte operationalisiert. Demgegenüber steht indes die theoretische Beschreibung des Konstrukts als komplexes, vielschichtiges und mehrdimensionales Phänomen. Dem damit vorliegenden Desiderat empirischer Spannungsmessung nimmt sich der vorliegende Beitrag an. Basierend auf einer umfangreichen Literatursichtung wird Spannung auf theoretischer Ebene definiert als subjektives Rezeptionserleben, welches sich durch die drei Merkmalsdimensionen a) Präferenz für einen Handlungsausgang, b) Unsicherheit und c) Risiko für den:die Protagonist:in konstituiert. Vor dem Hintergrund dieser Definition wird die *Skala Spannungserleben* (SSE) präsentiert, die im Rahmen zweier explorativer Vorstudien und einer Hauptstudie (EFA) zunächst entwickelt und in einer weiteren Hauptstudie konfirmatorisch getestet (CFA) sowie mithilfe eines nomologischen Netzwerks validiert wurde.

Schlagwörter: Spannung, Skala, Rezeptionserleben, Skalenentwicklung, Konstruktvalidierung.

Abstract: “Suspense is when it’s suspenseful.”¹ For many years this sentence coined the (German) empirical research on suspense. Post-receptive measurement instruments usually understood suspense as one-dimensional or as a subdimension of other constructs. This however stands in contrast to the theoretical description of suspense as a complex, multi-layered and multi-dimensional phenomenon. This paper addresses the desideratum of measuring suspense. Based on a detailed synopsis, suspense is described as subjective narrative experience which is characterized by three dimensions: a) preference for a certain outcome of a story, b) uncertainty, and c) risk perception. In view of the above the *Skala Spannungserleben* (SSE) [scale for measuring suspense (SSE)] is presented. The scale was developed, confirmed and validated on the basis of two pre- and two main studies. It was

1 This sentence is the analogous translation of the title of a paper published by Vorderer in 1994.

initially tested by exploratory factor analysis (EFA), followed by confirmatory factor analysis (CFA) and finally validated by a nomological network.

Keywords: Suspense, scale, narrative experience, scale development, construct validity.

1. Einleitung

Ob bei Horrorfilmen im Kinosaal, *Stranger Things* auf *Netflix*, beim sonntäglichen *Tatort* oder den Abenteuern von *Harry Potter*, aber auch bei einer Simpsons-Episode: Zuschauer:innen erleben bei der Film- oder Serienrezeption *Spannung* und lieben diesen Nervenkitzel. Negative emotionale Regungen, welche das *Spannungserleben* bisweilen mit sich bringen kann (z. B. Zillmann, 1996; 1980), stehen dem nicht im Weg. Im Gegenteil: Medieninhalte werden nicht zuletzt aufgrund des Strebens nach „Genuss, Abenteuer, Spannung, Abwechslung und [dem] Ausleben emotionaler Bedürfnisse“ (Weinacht & Hohlfeld, 2007, S. 186) selektiert und ausgewählt. Dabei führt Spannungserleben zu einem gesteigerten Rezeptionsgenuss (u. a. Thissen et al., 2021; Vorderer et al., 2004) und motiviert Rezipient:innen zur (weiteren) Rezeption (z. B. Czichon, 2019; Schulze, 2006). Dies wiederum kann Rezeptionsphänomene wie *Binge-Watching* erklären (Czichon, 2019), das aktuell vor allem auf Streaming-Plattformen, beispielsweise *Netflix* oder *Disney Plus*, praktiziert wird. Spannende Serien wie *Stranger Things* oder *Squid Game* sind in aller Munde. Spannung scheint ein Erfolgsgarant unterhaltender Medienerzeugnisse zu sein und lässt die Kassen klingeln.

Aus gesellschaftlicher und wirtschaftlicher Perspektive ergibt sich damit ein Phänomen von höchster Relevanz. Dieser Relevanz wird die wissenschaftliche Forschung jedoch nicht unbedingt gerecht, wie Hastall (2014, S. 268) konstatiert: „Angesichts dieser Bedeutsamkeit erscheint der theoretische und empirische Stand der Spannungsforschung unbefriedigend.“

Die aus kommunikationswissenschaftlicher Perspektive relevante Betrachtung von Spannung lässt sich in zwei Traditionen ordnen (Vorderer, 1994, S. 323f.): Zum einen werden *werkimmanente* Merkmale als spannungsgenerierend analysiert (dazu u. a. Wulff, 1993). Neben diesem – eher medienwissenschaftlichen – Zugang wird Spannung zum anderen aus Perspektive der Rezeptionsforschung als wichtiges Rezeptionsphänomen und als Komponente des *Unterhaltungserlebens* diskutiert (u. a. Czichon, 2019; Früh, 2002; Hoffner & Levine, 2005; Lewis et al., 2014; Oliver & Bartsch, 2010; Tamborini et al., 2011; Vorderer, 1994; Vorderer & Ritterfeld, 2009; Vorderer et al., 2004). Zentral für unsere Argumentation ist die zweite Tradition, da unser Erkenntnisinteresse vor allem dem Erleben von Spannung als kognitiver und affektiver Prozess der Rezipierenden gilt. Wir ziehen dennoch auch die Werksperspektive heran und stützen unsere Argumentation auch auf Erkenntnisse aus dieser Tradition.

In beiden Traditionen wird das Konstrukt *Spannung* als ein vielschichtiges, komplexes und mehrdimensionales Phänomen behandelt (u. a. Borringo, 1980; Carroll, 1990; Schulze, 2006; Zillmann, 1996). Trotz der umfangreichen theoretischen Diskussion verschiedener Facetten von Spannung wurde deren Messung in der Vergangenheit wenig elaboriert und eindimensional durchgeführt. Sie erfolgt

beispielsweise mit Fragen im Anschluss an die Rezeption, wie zutreffend Items wie „thrilling“ oder „exciting“ sind oder wie „gespannt“ man während der Rezeption war (Knobloch et al., 2004, S. 268f.; Knobloch-Westerwick et al., 2009, S. 755). Nur wenig weiter gehen die Autor:innen der Skala zur Erfassung des hedonischen und eudaimonischen *Unterhaltungserlebens* (Oliver & Bartsch, 2010; Schneider et al., 2019, S. 152). Hier wird Spannung mithilfe von vier Items als *eine* Subdimension des Unterhaltungserlebens operationalisiert. Gleiches gilt für die Skala zum Leseerleben von Appel et al. (2002; siehe auch Thissen et al., 2021), in der Spannung eine Subdimension von insgesamt 14 Dimensionen ist und mit fünf Items operationalisiert wird.

All diesen Operationalisierungen zugrunde liegt die konzeptionelle Vorstellung von Spannungserleben als postrezeptives Globalurteil. Mithin bezieht sich das Spannungserleben hier nicht auf ein Erleben auf der Mikro-Ebene, bezogen auf einzelne Sequenzen oder Szenen. Vielmehr wird es verstanden als Makro-Erleben, das sich handlungsübergreifend konstituiert (u. a. Brewer, 1996, S. 116; Pfister, 2001, S. 147). Von dieser Perspektive unbenommen ist, dass dieses Erleben auf der Makro-Ebene auch bereits begleitend zur Rezeption besteht und sich dabei in einer „dynamisierten Intratransaktion [...] permanent“ (Früh, 2002, S. 158) ändert. Allerdings bleibt im Anschluss an die Rezeption „die soeben anlässlich des Medienbeitrags konstruierte Makroemotion [...] noch eine Weile bestehen, und sie behält auch noch ihren konkreten episodischen Bezug zur Sendung“ (Früh, 2002, S. 172). Dieses nach der Rezeption weiterbestehende Erleben auf der Makro-Ebene verstehen wir hier als ein, unserer Konzeptualisierung und Operationalisierung von Spannungserleben zugrundeliegendes Globalurteil, das es durch unser Instrument zu messen gilt. Das Erleben entwickelt sich dabei kontinuierlich während der Rezeption und liegt anschließend in einer postrezeptiven Form mit handlungsübergreifendem Bezug vor.

Vor diesem Hintergrund kommt der vorliegende Aufsatz dem Desiderat hinsichtlich der empirischen Erfassung von Spannung nach. Ziel ist es, eine Skala zu entwickeln, die der Multidimensionalität des Konstrukts gerecht wird und zudem Anschlussfähigkeit an zahlreiche theoretische Modellierungen von Spannung aufweist. Daher werden ausgehend von einer systematischen Analyse der Spannungsliteratur die zentralen Dimensionen des Konstrukts sowie dessen Voraussetzungen und Konsequenzen abgeleitet. Auf dieser Basis wird das Konstrukt Spannungserleben definiert. Darauf aufbauend wird ein Messinstrument entwickelt und validiert. Zuletzt werden Befunde der Validierung vorgestellt und diskutiert. Das Fazit schließt mit Implikationen für die Verwendung der Skala bzw. einem Ausblick auf weitere Forschungsnotwendigkeiten.

2. Das Konstrukt Spannungserleben

Je nach Ansatz und Definition werden dem Konstrukt Spannungserleben unterschiedliche Auslöser und Facetten zugeordnet. Wie in Tabelle 1 in zusammengefasster Form dargestellt und später noch ausführlicher erläutert, betonen beispielsweise Carroll (1990), de Wied (1991), Öhding (1998) oder Zillmann (1996) die Einstellung der Rezipierenden zum:zur Protagonist:in. Appel et al. (2002,

S. 151) verstehen Spannung als eine „identifikatorische oder empathische emotionale Reaktion“ und heben damit die affektive Erlebenskomponente des Phänomens hervor. Diese Facette ist auch bei Borringo (1980), de Wied (1991), Eder (2007), Früh (2002), Fuchs (2000), Hant (1999), Öhding (1998) oder Zillmann (1996) zu finden. Wiederum zentral ist für Carroll (1990), Eder (2007) oder Zillmann (1996) die Antizipation des Handlungsausgangs. Borringo (1980), Carroll (1990), de Wied (1991), Früh (2002), Fuchs (2000), Hant (1999), Öhding (1998) oder Zillmann (1996) schreiben der Unsicherheit darüber, wie eine Geschichte zu Ende geht eine tragende Rolle bei der Konstituierung von Spannungserleben zu. Für Borringo (1980), Eder (2007), Hant (1999), Öhding (1998) oder Zillmann (1996) ist die Präferenz für einen bestimmten Handlungsausgang ein weiteres konstituierendes Merkmal. Und schließlich ist die Wahrnehmung eines großen Risikos für den:die Protagonist:in beispielsweise bei Borringo (1980), Carroll (1990), de Wied (1991), Hant (1999) oder Zillmann (1996) ein Auslöser für Spannungserleben.

Tabelle 1. Merkmale und Aspekte von Spannung in der Forschungsliteratur

	<i>Borringo,</i> <i>1980</i>	<i>Carroll,</i> <i>1990</i>	<i>de</i> <i>Wied,</i> <i>1991</i>	<i>Eder,</i> <i>2007</i>	<i>Früh,</i> <i>2002</i>	<i>Fuchs,</i> <i>2000</i>	<i>Hant,</i> <i>1999</i>	<i>Öh-</i> <i>ding,</i> <i>1998</i>	<i>Zill-</i> <i>mann,</i> <i>1996</i>
Einstellung zu Protagonist:in		X	X					X	X
Empathie/Identifikation	X		X	X	X	X	X	X	X
Antizipation von Handlungs- ausgängen		X		X					X
Unsicherheit über Handlungs- ausgang	X	X	X		X	X	X	X	X
Risikowahrnehmung	X	X	X				X		X
Präferenz für Handlungs- ausgang	X			X			X	X	X

Wir gehen davon aus, dass sich die in den verschiedenen theoretischen Zugängen genannten Merkmale von Spannung analytisch in notwendige und hinreichende Bedingungen für Spannungserleben unterscheiden lassen. Zu den notwendigen Bedingungen zählen Voraussetzungen wie die Einstellung zum:zur Protagonist:in oder empathische Prozesse. Diese Voraussetzungen müssen gegeben sein, damit Spannung überhaupt eintreten kann, aber sie selbst konstituieren noch nicht das Erleben von Spannung. Dies ist Aufgabe der hinreichenden Bedingungen. Hierbei handelt es sich, wie wir im Folgenden zeigen werden, um (1) eine Präferenz für einen Handlungsausgang, (2) Unsicherheit über den Verlauf und (3) ein wahrgenommenes Wagnis und damit einhergehendes Risiko für den:die Protagonist:in. Schließlich hat Spannungserleben auch weitere affektive Erlebensphänomene zur Folge.

2.1 Voraussetzungen für Spannungserleben

Um sich dem Kern des *Spannungserlebens* zu nähern, soll zunächst ein keineswegs trivialer Aspekt in den Vordergrund gerückt werden: Narrationen sind getragen von Personen und Figuren (Früh & Frey, 2014, S. 327; Mikos, 2008, S. 19; Wünsch, 2014, S. 223). Diese werden werkseitig in den Fokus der Narration gerückt. Ihre Ziele, Handlungen oder affektiven Zustände werden Rezipient:innen also fortwährend präsentiert – und dienen als Grundlage für deren *Unterhaltungserleben* (Busselle & Bilandzic, 2008). Hierfür vollziehen Rezipient:innen einen *Deictic Shift* (Segal, 1995). Damit verorten sie sich selbst in der Welt der Geschichte, können diese *immersiv*, d. h. *narrativ erleben* (Busselle & Bilandzic, 2009) sowie im Rahmen *empathischer* (Leibetseder et al., 2001; Shen, 2010) oder *identifikatorischer* (Cohen, 2001) Prozesse die Perspektive der handelnden Figuren einnehmen. Der wohl wichtigste Unterschied zwischen Identifikation und Empathie ist, dass bei ersterem „das Bewusstsein um die eigene Identität und die mediale Vermitteltheit der übernommenen Identität in den Hintergrund rückt“ (Wünsch, 2014, S. 234) und daher vermutet werden kann, dass dadurch die Perspektivübernahme durch die Zuschauer:innen intensiver erlebt wird. In den genannten Theorien werden die beiden Phänomene allerdings nur selten trennscharf verwendet, sodass wir für die vorliegende Betrachtung Identifikation vereinfacht als eine besonders intensive Form von Empathie verstehen (Wünsch, 2014) oder wie auch Shen (2010, S. 507) als „associative empathy“.

In vielen Spannungstheorien wird diese Einnahme der Perspektive handelnder Figuren bereits als erstes zentrales Definitionsmerkmal von Spannungserleben gewertet (Borrino, 1980; de Wied, 1991; Eder, 2007; Früh, 2002; Fuchs, 2000; Hant, 1999; Öhding, 1998; Zillmann, 1996): Je stärker sich die Rezipient:innen mit den Protagonist:innen identifizieren, desto stärker falle auch das Spannungserleben aus. Borrino (1980, S. 46) begründet dies mit der Übernahme zentraler Gefühle der handelnden Figuren durch die Rezipient:innen. Diese Auffassung teilt Früh (2002, S. 190): „Durch die Identifikation mit dem Protagonisten agiert der Zuschauer in der virtuell angenommenen Rolle innerhalb der dargestellten virtuellen Welt.“ Im Gegensatz zu Borrino (1980) weist Früh (2002, S. 190) aber darauf hin, dass dabei ein gewisses Restbewusstsein der medialen Vermitteltheit vor-

handen bleibt: Den Rezipient:innen sei grundsätzlich latent bewusst, dass es sich lediglich um ein Spiel, eine virtuelle Realität, handelt.

Grundsätzlich empfinden Rezipient:innen aber nicht mit jeder Person in gleichem Umfang mit. Stattdessen spielt die Einstellung zur Hauptfigur ebenfalls eine Rolle und Empathie wird vor allem jenen Figuren zuteil, die in ihrem Verhalten sympathisch scheinen und die den moralischen Anforderungen des Publikums genügen. Diesen wird eine positive Einstellung – eine positive *affektive Disposition* – gegenüber entwickelt, wie Zillmann (1980, S. 135; 1996, S. 202) in seiner Definition von Spannung hervorhebt. Figuren, die unsympathisch sind und moralisch fragwürdig agieren (siehe auch Carroll, 1996, S. 78ff.), sog. Antagonist:innen, lehnen Rezipient:innen ab, wobei es in der Beurteilung hinsichtlich der moralischen Korrektheit zu interindividuellen bzw. situativ gar intraindividuellen Unterschieden kommen kann. In diesem Sinne verstehen wir hier Einstellungen zu Figuren als Bewertungen der Rezipient:innen, die sich in ihrer Valenz (sympathisch oder unsympathisch, gut oder schlecht, mögen oder nicht mögen) und ihrer Stärke unterscheiden (Mahrt, 2014, S. 133).

Damit lässt sich an dieser Stelle festhalten, dass es sich sowohl bei einer positiven *affektiven Disposition* als auch den Prozessen *Empathie* und *Identifikation* zweifelsohne um notwendige Bedingungen von Spannungserleben handelt, die aber nicht hinreichend sind. Nur wer mit einer Hauptfigur *mitfühlen* kann, wird am Ende mit ihr *mitfiebern* können. Umgekehrt könnte man auch sagen, eine Person, der die Hauptfigur egal ist, wird keine Spannung verspüren. Doch das empathische Miterleben mit einer unsympathischen Figur oder die Identifikation mit ihr alleine machen noch kein Spannungserleben: Empathisches Miterleben könnte auch lediglich als Begleiterscheinung *parasozialer Interaktion* (u. a. Hartmann et al., 2004; Horton & Wohl, 1956; Giles, 2002), *Transportation* (Green & Brock, 2000, 2002; Green et al., 2004) oder vergleichbarer Phänomene auftreten.

Ähnlich verhält es sich mit der Antizipation des Handlungsverlaufs, die in einigen Spannungstheorien als zweites zentrales Kriterium präsentiert wird (Carroll, 1990; Eder, 2007; Zillmann, 1996): Narrationen zeichnen sich dadurch aus, dass sie zu Beginn eine Reihe von Fragen aufwerfen, die erst im weiteren Verlauf der Narration beantwortet werden. Zur Verarbeitung der Narration antizipieren Rezipient:innen den potenziellen Handlungsverlauf. Die Rezipient:innen generieren (unbewusst) Vermutungen und erwarten zudem, dass die nachfolgenden Szenen Antworten auf ihre Fragen geben (Carroll, 1990, S. 132f.; Schwan, 2014, S. 191). Damit gelingt es, die Narration überhaupt kognitiv verfügbar zu machen; sie zu verstehen (Kintsch, 1993). Es handelt sich also auch hier um eine notwendige, nicht jedoch hinreichende Bedingung für Spannungserleben: Das reine Vorausahnen einer Handlung macht ebenfalls noch keine Spannung.

2.2 Konstituierende Merkmale von Spannungserleben

Präferenz für einen Handlungsausgang

Wichtiger als das bloße Vorausnehmen eines bestimmten Handlungsausgangs ist die *Präferenz* für einen bestimmten Ausgang der Handlung (Borrino, 1980; Carroll, 1990, 1996; Eder, 2007; Hant, 1999; Öhding, 1998; Zillmann, 1996).

Spannung ist laut de Wied (1991), de Wied et al. (1994) und Zillmann (1980, 1996) eine negative Stressreaktion, die dadurch entsteht, dass den Rezipierenden das Schicksal der handlungstragenden Figur nicht gleichgültig ist. Gehofft wird auf einen positiven Ausgang für die gemochte Figur. Gefürchtet ist hingegen ein Unheil für diese (Zillmann, 2004). Neben dem gewünschten positiven Handlungsausgang, existiert also ein möglicher, alternativer, negativer Handlungsausgang für die Hauptfigur. Durch diese drohende Gefahr entsteht beim Publikum empathischer *Dysstress* (Zillmann, 1991, S. 139; siehe auch de Wied et al., 1994). Dies geschieht aber nur, solange das Publikum der Figur wohlgesinnt ist. Andernfalls entsteht ein Gefühl, das Zillmann (1996, S. 217) als *Counter-Empathie* bezeichnet und eine Art Schadenfreude darüber, dass eine wenig gemochte Figur nicht erfolgreich war.

Was lässt sich daraus für das Spannungserleben ableiten? Wem die Hauptfigur und damit das Schicksal dieser Figur egal ist, wird keine Spannung verspüren. Eine Präferenz für einen bestimmten, für die Hauptfigur guten Handlungsverlauf zu haben, ist damit eine notwendige und – gemeinsam mit den beiden folgenden Dimensionen – hinreichende Bedingung für das individuelle Spannungserleben. Eine klare Handlungspräferenz manifestiert sich in dem Wunsch, dass sich die Hauptfigur am Ende aus einer bedrohlichen Lage befreien kann und/oder ihr Ziel erreicht. Gleichzeitig zeigt sie sich in der Befürchtung um ihr Scheitern.

Unsicherheit

Während der Rezeption antizipieren Zuschauer:innen verschiedene Handlungsausgänge, deren Realisation für sie ungewiss ist. Damit haben sie allenfalls eine Vermutung, wie die Situation *tatsächlich* ausgeht. Dies verursacht *Unsicherheit* – einen Zustand, der u. a. Carroll (1990) zufolge zentral für das Spannungserleben ist. Während der Rezeption spannender Inhalte befinden sich die Rezipient:innen damit fortwährend in Erwartung darüber, welche Gefahren und Risiken als nächstes auf die handlungstragenden Figuren zukommen werden (Borrino, 1980, S. 46) (siehe auch „Risikowahrnehmung“).

Nach Früh (2002, S. 189) kommt es im Individuum durch diese Konfrontation mit „Ungewissheit, Unsicherheit, Unvorhersehbarkeit, Widersprüchlichkeit etc.“ zu einem externalen (aktiven oder passiven) Kontrollverlust. Diese Dysbalance zwischen dem Individuum und seiner Umwelt führt zu einem Verlust der planvollen und gezielten Steuerung der „eigenen Kognitionen, Emotionen und Verhaltensweisen und sonstigen physischen und physiologischen Reaktionen“ (Früh, 2002, S. 190).

Als wichtige Voraussetzung betont Carroll (1990, S. 137f.), dass es wenig wahrscheinlich erscheinen darf, dass die Geschichte ein gutes Ende nehmen wird. Zillmann (1980, S. 138f.; 1996, S. 200) konkretisiert diesbezüglich, dass es sich bei Unsicherheit nicht um eine sich linear verhaltende Bedingung handelt: „Uncertainty is thus at a maximum when the odds for a desired or a feared outcome are 50–50“ (Zillmann, 1996, S. 200).

Was also passiert, wenn sich die Rezipient:innen zu einem bestimmten Punkt der Geschichte zu 100 Prozent sicher sind, dass es zu einem schlechten Ende für den:die Protagonist:in kommt? Dieses Szenario würde das fortwährende Hoffen und Bangen überflüssig machen. Die Folge wäre nicht mehr länger Spannung, sondern Enttäuschung und Traurigkeit (Zillmann, 1980, S. 134). Es muss also trotz des Pessimismus auf Seiten der Zuschauer:innen ein ‚Fünkchen‘ Resthoffnung bestehen bleiben, dass doch alles gut ausgeht. Ebenso andersherum, wenn der Ausgang relativ sicher positiv sein wird, muss ein ‚Fünkchen‘ Zweifel bestehen, ob es tatsächlich gut ausgehen wird.

Hervorzuheben sei an dieser Stelle, dass der Grad der Unsicherheit nicht ausschließlich aus der Sicht der Rezipient:innen bewertet wird. Wie sonst ließe sich erklären, dass Rezipient:innen bei der wiederholten Rezeption bereits rezipierter (identischer) Medieninhalte erneut Spannung erleben können? Dieses von Carroll (1996) als *Paradox of Suspense* beschriebene Phänomen wird aufgelöst durch die Vergegenwärtigung der Phänomene, die im vorstehenden Abschnitt als Voraussetzungen für Spannungserleben beschrieben wurden: Durch den *Deictic Shift* befindet sich das Publikum mitten in der Geschichte. Damit erleben sie die Handlung aus der Sicht des:der Protagonist:in und sind somit auch in der Lage, sich aufgrund verschiedener empathischer oder identifikatorischer Prozesse in dessen:deren Unsicherheit hinsichtlich des Handlungsverlaufs hineinzuversetzen.

Risikowahrnehmung

In der Literatur findet sich vielfach die Vorstellung, dass die Wahrnehmung eines Wagnisses bzw. eines Risikos für das Wohlergehen der Hauptfigur ein wichtiger Bestandteil von Spannungserleben ist. Zillmann (1996, S. 208) definiert Spannung beispielsweise als „a noxious affective reaction that characteristically derives from the respondents’ acute, fearful apprehension about deplorable events that threaten liked protagonists.“ Dass die Geschichte eine:n Bösewicht:in enthält, der:die das Wohlergehen eine:r Protagonist:in glaubwürdig gefährdet, setzt er dabei als zwingend notwendig voraus. Dem schließt sich de Wied (1991, S. 16) an, indem er an, als dass Spannung nur dann entstehen könne, wenn die Narration Hinweise enthält, dass der handlungstragenden Figur etwas Schlimmes, zumindest aber etwas Unheilvolles widerfahren kann. Je höher auf Seiten der Rezipierenden die Risikowahrnehmung für die Hauptfigur ist, desto höher ist das Spannungserleben (Borrino 1980, S. 53; Carroll 1990, S. 137f.; de Wied, 1991, S. 16). Damit wird „von der Prämisse ausgegangen, dass ein Geschehen, dessen Entwicklung ohne Bedeutung für jedermann wäre, keine Spannung erzeugen würde“ (Schulze, 2006, S. 22). Dabei muss es nicht zwangsläufig – wie von Zillmann (1996, S. 209) gefordert – eine:n Bösewicht:in geben, der:die das Wohlergehen eines:einer

Protagonist:in gefährdet. Stattdessen muss lediglich ein aus Sicht des Publikums glaubwürdiges Risiko bestehen und sei es das Risiko, einen bestimmten Gewinn nicht zu erhalten.

Daraus ergibt sich auch: Es muss nicht immer gleich das Leben der handlungs-tragenden Figur bedroht sein. In jedem Falle aber muss viel für den:die Protagonist:in auf dem Spiel stehen, d. h., es muss sich um etwas für die handlungstragende Figur Relevantes handeln. Es kann bspw. auch ‚nur‘ das Liebesglück der Hauptfigur in Gefahr sein. Solange die Handlung für den Ausgang der Geschichte von Bedeutung ist, kann Spannung entstehen.

Fehlen dem Publikum hingegen die Informationen für eine Risikoabschätzung und ist es sich folglich der Relevanz der Situation für das Wohlergehen der Figur nicht bewusst, dann ist auch das Rezeptionserleben ein ganz anderes: Es wird keine Spannung verspüren. Stattdessen dürfte es nun ziemlich überrascht sein, ob einer plötzlichen Wendung. Überraschung ist ein anderes Rezeptionsphänomen als Spannung (Brewer, 1996). Die Wahrnehmung eines Risikos für die Protagonist:in trägt mithin erst dann zur Spannung bei, wenn dieses zwar absehbar aber keineswegs sicher ist (siehe auch „Unsicherheit“).

Spannung konstituiert sich der Argumentation zufolge damit aus den drei notwendigen und hinreichenden Dimensionen: (1) eine klare Präferenz für einen Handlungsausgang, (2) Unsicherheit über den Handlungsausgang, verbunden mit (3) der Wahrnehmung eines bedeutungsvollen Risikos für den:die Protagonist:in.

2.3 Konsequenzen von Spannungserleben

Spannungserleben wird häufig als eine Subdimension von *Unterhaltungserleben* konzeptualisiert (z. B. Czichon, 2019; Lewis et al., 2014; Oliver & Bartsch, 2010; Tamborini et al., 2011; Vorderer & Ritterfeld, 2009; Vorderer et al., 2004). Spannung trägt somit zum *Rezeptionsgenuss* bzw. *Enjoyment* bei. Davon abhängig ist zudem auch die *Gesamtbewertung* eines Stimulus als unterhaltend. Der Zusammenhang zwischen Spannung und Unterhaltung ist empirisch bereits gut bestätigt (im Überblick Hastall, 2014).

Während Rezeptionsgenuss bzw. *Enjoyment* zu einer positiven Bewertung einer Narration führt, hat die positive Bewertung wiederum *motivationale Effekte* und führt beispielsweise dazu, die *Rezeption fortzuführen* (Czichon, 2019; Schulze, 2006). Die Motivation, eine Rezeption zu beenden oder sich einem anderen Inhalt bzw. einer anderen Tätigkeit zuzuwenden, sollte demnach im Gegenzug bei hohem Enjoyment gering ausfallen. Gleiches gilt für *Nebentätigkeiten* bei der Rezeption. Ist der Rezeptionsgenuss hoch, werden Nebentätigkeiten eher unterlassen. Schließlich kann die motivationale Wirkung auch über die entsprechende Rezeptionssituation hinausgehen: So werden positiv bewertete Situationen eher wieder aufgesucht. Das bedeutet, ein postrezeptiv positiv bewerteter Stimulus wird mit größerer Wahrscheinlichkeit *wiederholt* selektiert und rezipiert. Nicht zuletzt werden Personen, nachdem sie Spannungserleben empfunden haben, diese Rezeption auch selbst als „spannend“ bewerten. Denn, wenn auch unsere theoretische Konzeptualisierung von Spannung systematisch aus der Forschungsliteratur abgeleitet

wurde, bleibt sie dennoch weitgehend konsistent zum *Alltagsverständnis* des Phänomens Spannung.

3. Entwicklung der Skala Spannungserleben (SSE)

In den bisherigen Ausführungen wurde *Spannungserleben* als ein rezeptionsbegleitendes, subjektives Erleben definiert, das durch drei Merkmale gekennzeichnet ist: (1) Die Rezipierenden weisen eine Präferenz für einen bestimmten Handlungsausgang für die Protagonist:innen auf. (2) Die Rezipierenden sind unsicher hinsichtlich dieses Handlungsausgangs. (3) Die Rezipierenden nehmen wahr, dass für die Protagonist:innen ein Risiko für ihr Wohlergehen besteht.

In einem nächsten Schritt soll nun ein Messinstrument, eine Skala zur Erfassung dieses so definierten Konstrukts, entwickelt und einer Überprüfung unterzogen werden. Um das von uns definierte Konstrukt valide mit der zu entwickelnden Skala zur Erfassung des Spannungserlebens (SSE) erfassen zu können, muss diese (1) der *Mehrdimensionalität* des Konstrukts gerecht werden und somit alle drei Dimensionen angemessen erfassen. Hinsichtlich dieser drei Dimensionen muss die Skala (2) *das subjektive Erleben* während der Rezeption erfassen und dabei zudem (3) auf das Erleben während der *gesamten* Rezeptionsepisode zielen und nicht bspw. ausschließlich auf prägnante Szenen, den Schluss oder einen Cliffhanger. Eine naheliegende rezeptionsbegleitende Erfassung mithilfe einer Befragungsskala würde mit einer starken Reaktivität des Messinstrumentes einhergehen, da der Rezeptionsprozess durch Unterbrechungen zu stark gestört würde. Um dieses Problem zu vermeiden und in dieser Hinsicht eine reliable Messung zu ermöglichen, soll (4) eine Skala zum postrezeptiven Einsatz, unmittelbar im Anschluss an die Rezeption entwickelt werden. Es erfolgt also eine Messung des Rezeptionserlebens in seiner postrezeptiven Ausformung, das sich zwar während der vorangehenden Rezeption kontinuierlich entwickelt hat, aber seinen „konkreten episodischen Bezug zur Sendung“ (Früh, 2002, S. 172) noch nicht verloren hat. Leitend bei der Skalenentwicklung war weiterhin das Ziel, ein forschungsökonomisch effizientes Instrument zu entwickeln, das sich (5) flexibel sowohl im Feld als auch im Labor und zudem in Selbstadministration bearbeiten lässt. Für das von uns definierte Konstrukt des Spannungserlebens gehen wir zwar davon aus, dass es sich bei der Rezeption unterschiedlicher Medien nicht kategorial unterscheidet. Eine Befragungsskala, die gleichermaßen für alle Medien (Film, Roman, Hörspiel etc.) einsetzbar ist, könnte in ihren Formulierungen hingegen nur in einer abstrakten Form Bezug auf die vorangegangene Rezeption nehmen. Eine solche, von den Untersuchungsteilnehmer:innen dann geforderte hohe Abstraktionsleistung birgt die Gefahr einer weniger validen und reliablen Messung (z. B. Möhring & Schlütz, 2013, S. 186f.) Daher fokussieren wir uns (6) zunächst ausschließlich auf das Rezeptionserleben audiovisueller Angebote. Aus dem gleichen Grund soll sich (7) die Skala zunächst ausschließlich auf die Rezeption abgeschlossener Narration beziehen.

Zur empirischen Erhebung des Konstrukts Spannungserleben schlagen wir daher eine postrezeptive schriftliche Befragung vor. Gemessen werden dabei die drei Dimensionen *Präferenz für einen Handlungsausgang* (z. B. „Ich wollte, dass die

Hauptfigur unbedingt ihre Ziele erreicht.“), *Unsicherheit* (z. B. „Die Hauptfigur konnte sich lange Zeit nicht sicher sein, ob sie am Ende erfolgreich sein würde.“) und *Risikowahrnehmung* (z. B. „Für die Hauptfigur stand viel auf dem Spiel.“).

Zur Entwicklung und Validierung der SSE wurden zwei Hauptstudien sowie zwei kleinere, explorative Vorstudien durchgeführt. Ausgangspunkt bildete ein Itempool von ca. 50 Items, der auf Basis theoretischer Vorüberlegungen und unter Berücksichtigung der vorstehenden Literatur sowie der aufgeführten Anforderungen an das Messinstrument generiert wurde. Dieser erste Itempool wurde in zwei Vorstudien ($n_1 = 71$, $n_2 = 165$) explorativ getestet, um erste Hinweise zur Eignung einzelner Items zu erhalten. Im Anschluss wurde in einer ersten kombinierten Feld- ($n = 99$) und Laborstudie ($n = 123$) eine systematische Testung von 30 Items vorgenommen. Dabei wurde u. a. eine *explorative Faktorenanalyse* sowie eine *Reliabilitätsprüfung* durchgeführt. Auf Basis dieser Ergebnisse erfolgte eine erneute Reduktion auf die vier geeignetsten Items pro Dimension, die einer ersten *konfirmatorischen Faktorenanalyse* unterzogen wurden. Im dritten Schritt wurde mit den Daten einer erneuten Laborstudie ($n = 226$) eine weitere konfirmatorische Faktorenanalyse durchgeführt (Studie 2). Die *Validierung* der Skala erfolgte mittels eines *nomologischen Netzwerks* (Feigl, 1958). Die dazu eingesetzten Validierungskonstrukte wurden in der ersten und zweiten Studie miterhoben. Diese werden in Abschnitt 7.2 vorgestellt.

4. Explorative Vorstudien zur Itemgenerierung

Der auf Basis der theoretischen Vorüberlegungen und unter Berücksichtigung der aufgeführten Anforderungen entstandene Itempool wurde in zwei explorativen Vorstudien einer ersten Inspektion unterzogen. Ziel der ersten Itemgenerierung war es, die drei in unserer Definition genannten konstituierenden Merkmale von Spannungserleben abzubilden: *Präferenz für einen Handlungsausgang*, *Unsicherheit* und *Risikowahrnehmung*. Dabei wurde auf eine ausgewogene Itemschwierigkeit sowie die Verwendung sowohl positiv als auch negativ formulierter Aussagen geachtet. Diese Items wurden den Proband:innen direkt im Anschluss an die Rezeption eines audiovisuellen, narrativen Angebots (z. B. Spielfilm) vorgelegt. Anhand einer siebenstufigen Ratingskala (von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“) sollten diese dann angeben, wie stark diese Aussagen retrospektiv in Bezug auf das Erleben während der Rezeption zutreffen („Im Folgenden interessiert uns, wie Sie sich während des Films gefühlt haben, wie Sie die Geschichte des Films empfunden haben und was Ihnen während des Sehens durch den Kopf ging.“)². Da sich zahlreiche Items auf das Erleben hinsichtlich des:der Protagonist:in beziehen, wurden die Befragten außerdem zunächst aufgefordert, ihre:n „Filmheldin oder Filmhelden“ zu benennen („In einigen der folgenden Fragen wird es um die ‚Hauptfigur‘ des Films gehen. Wir meinen damit die für die Handlung wichtigste, Ihnen aber auch sympathische Figur – sozusagen die Film-

2 Der komplette Wortlaut der Instruktion ist im Anhang des Beitrags abgedruckt.

heldin oder den Filmhelden. [...] Können Sie uns bitte als Erstes aufschreiben, wer in dem Film eben die Hauptfigur war?“).

Die Eignung dieses Vorgehens und des Itempools wurde in zwei explorativen Erhebungen getestet. Zunächst wurden insgesamt 30 potenzielle Items einer anfallenden Stichprobe ($n = 71$; Durchschnittsalter 36 Jahre, 64 % weiblich, 41 % Studierende) und in einer zweiten Erhebung 13 Items einer weiteren, ebenfalls anfallenden Stichprobe ($n = 165$; Durchschnittsalter 31 Jahre, 62 % weiblich, 54 % Studierende) vorgelegt. Diese 30 bzw. 13 Items wurden von den Befragten im alltäglichen, häuslichen Kontext, in einer selbstbestimmten Rezeptionssituation jeweils direkt im Anschluss an die Rezeption eines selbstgewählten Spielfilms (Vorstudie 1 und 2) oder einer selbstgewählten Episodenserie (nur Vorstudie 2) schriftlich beantwortet. In der Vorstudie 2 wurden zudem zehn der 71 Befragungen durch eine teilnehmende Beobachtung in Form der *Methode des lauten Denkens* begleitet, um weitere Erkenntnisse in Bezug auf das Itemverständnis zu generieren.

Auf Basis des Feedbacks zur Verständlichkeit der Itemformulierungen im Rahmen des lauten Denkens (z. B. Bilandzic, 2005) sowie der Analyse der psychometrischen Kennwerte wurden zunächst unpassende Items aussortiert bzw., soweit möglich, einer Überarbeitung unterzogen. Die danach verbliebenen Items der Vorstudien wurden in einem zweiten Schritt mehreren explorativen Faktorenanalysen unterzogen. Außerdem wurde deren diskriminante Validität zu verwandten Konstrukten auf Basis einfacher korrelativer Analysen untersucht. Im Ergebnis lieferte dieser zweite Analyseschritt keine stimmige Gesamtlösung zu einer Faktorstruktur zur Messung der von uns postulierten Dimensionen von Spannungserleben. Im Resultat führte dies zu einer *kompletten Itemrevision*, sodass letztlich keines dieser Items im Wortlaut in die finale Skala Eingang fand. Dennoch lieferten diese Vorstudien in ihrer Gesamtschau wichtige Erkenntnisse zur Operationalisierung des Spannungserlebens:

Die Dimension *Präferenz für einen Handlungsausgang* wurde in den Vorstudien u. a. auch mithilfe von Formulierungen operationalisiert, die den Handlungsausgang retrospektiv beurteilten (z. B. „Ich habe mich am Ende für die Filmfigur gefreut.“). Dieser Zugang erwies sich in unserer Evaluation einer auf das prospektive Erleben in der Rezeptionsphase bezogenen Formulierung (z. B. „Ich habe gehofft, dass es ein gutes Ende für die Filmfigur gibt.“) als unterlegen.

Hinsichtlich der Dimension *Unsicherheit* erwies sich eine Formulierung, die auch die eigene Unsicherheit (des:der Rezipient:in) adressierte (z. B. „Ich hatte keine Ahnung, wie der Film ausgehen würde.“), als nicht zielführend. Formulierungen, die hingegen die Einschätzung der Unsicherheit innerhalb der Story, für den:die Protagonist:in adressierten (z. B. „Das Schicksal der Hauptfigur hing in der Schwebe.“), schnitten in der Evaluation besser ab. Zudem zeigte sich im Rahmen der *Methode des lauten Denkens*, dass sich die Antworten der Befragten nicht immer gleichermaßen – wie von uns angenommen – auf Handlungsverläufe bezogen, die für das Wohlergehen und die Zielerreichung der Hauptfigur von hoher Relevanz sind (so wie in unserer Definition von Spannungserleben gefordert). Es wurde vielmehr auch Unsicherheit hinsichtlich dramaturgischer oder filmischer

Gestaltungsmittel und teilweise auch nur bezogen auf einzelne Szenen oder gar Sequenzen adressiert.

5. Studie eins: Explorative Faktorenanalyse und Reliabilitätsprüfung

Auf Basis dieser Erkenntnisse entstand ein revidierter Itempool von insgesamt 31 Items zur Messung der drei Dimensionen von *Spannungserleben*, der im Zuge einer gemeinsamen Labor- und Feldstudie reduziert und evaluiert wurde. Außerdem wurden weitere mit dem *Spannungserleben* verknüpfte Konstrukte mithilfe einschlägiger Skalen erfasst, um Hinweise hinsichtlich der *diskriminanten Validität* zu erhalten (vgl. Abschnitt 7).

5.1 Methodik

Zur Überprüfung des revidierten Itempools und mithin zur Bestimmung der Dimensionalität der Skala sowie der Reduktion der Items wurde der Itempool einer weiteren Stichprobe vorgelegt. Diese Daten wurden dann einer *explorativen* und einer ersten *konfirmatorischen Faktorenanalyse* sowie damit einhergehend einer *Reliabilitätsprüfung* unterzogen. Insgesamt nahmen an dieser ersten Studie 222 Proband:innen teil. Sie wurde in Form einer gemischten Feld- ($n = 99$) und Laborstudie ($n = 123$) umgesetzt. In beiden Fällen handelte es sich um eine anfallende Stichprobe (Feldstudie: Durchschnittsalter 26 Jahre, 72 % weiblich; Laborstudie: Durchschnittsalter 21 Jahre, 79 % weiblich).

Im Zuge der *Feldstudie* wurden die 99 Teilnehmenden gebeten, unmittelbar nach der Rezeption einer Serie oder eines Films den Fragebogen auszufüllen. Die Entscheidung, welchen Stimulus sich die Proband:innen ansahen, wurde dabei insofern leicht eingeschränkt, als dass für den Feldzeitraum eine Liste mit ca. 50 ausgewählten Sendungen im *Free-TV* erstellt wurde. Aus den in diesem Programm gelisteten Sendungen sollten die Proband:innen einen für sie interessanten Medieninhalt selektieren. Durch dieses Vorgehen sollte einerseits sichergestellt werden, dass tatsächlich nur geeignete narrative Inhalte gewählt wurden und zum anderen auch die Varianz hinsichtlich der genutzten Genres gering genug ausfiel, um diese in Bezug auf das gemessene Spannungserleben einem Gruppenvergleich unterziehen zu können. Der Fragebogen enthielt neben den 31 Items des revidierten Itempools weitere Konstrukte zur späteren Validierung der Skala (siehe Abschnitt 7). Die Instruktion der Befragten zur Bearbeitung sowie die verwendete siebenstufige Ratingskala stammten aus den bereits vorgestellten Vorstudien.

Die *Laborstudie* wurde im Rahmen verschiedener einführender kommunikationswissenschaftlicher Lehrveranstaltungen durchgeführt. Dabei kamen jeweils verschiedene Stimuli zum Einsatz, bei denen von einem möglichst unterschiedlichen Spannungspotential ausgegangen werden konnte. So wurde einem Teil der Teilnehmenden eine Episode der TV-Serie *Friends* ($n = 73$) gezeigt, einem anderen eine Episode der *Simpsons* ($n = 35$) und einem letzten Teil eine Episode der Serie *Grey's Anatomy* ($n = 15$). Im Anschluss an die Rezeption der Stimuli erhielten die Teilnehmenden der Laborstudie den gleichen Fragebogen wie die Teilnehmenden

der Feldstudie, ergänzt um einige Items zur Erfassung weiterer Konstrukte, die später dann der Validierung der Skala dienen sollten (siehe Abschnitt 7).

5.2 Itemanalyse und -reduktion

In einem ersten Analyseschritt wurden die deskriptiven Verteilungs- und Lageparameter sowie weitere psychometrische Kennwerte aller 31 Items einer Inspektion unterzogen (Tab. 2). Diese wies keine signifikanten Auffälligkeiten auf, die zum Ausschluss von Items geführt hätten.

In einem anschließenden explorativen, iterativen Prozess wurde der Itempool zahlreichen explorativen Faktorenanalysen unterzogen und dabei schrittweise die Anzahl der berücksichtigten Items reduziert. Die Entscheidung, welche Items aus der Analyse ausgeschlossen wurden, orientierte sich dabei an schlechten Hauptladungen auf den zugehörigen Faktor, hohen Nebenladungen und geringen Kommunalitäten sowie in der letzten Phase dieses Prozesses an Hinweisen auf reliabilitäts erhöhende Itemausschlüsse im Rahmen von Reliabilitätsanalysen (*McDonald's Omega* (vgl. Hayes & Coutts, 2020)).

Tabelle 2. Psychometrische Kennwerte des in Studie 1 verwendeten Itempools

	M	SD	Schiefe	Kurtosis
Ich habe mir ein bestimmtes Ende gewünscht.	4.31	2.06	-0.24	-1.31
Ich hatte eine Vorliebe für einen bestimmten Ausgang des Films.	4.49	2.10	-0.47	-1.16
Ich habe mir ein gutes Ende für die Hauptfigur gewünscht.	5.30	1.54	-0.81	0.07
Ich hätte jeden Ausgang des Films gut gefunden.	2.85	1.84	0.76	-0.55
Ich hatte keine bestimmte Vorliebe, was das Ende des Films anging.	3.55	2.12	0.33	-1.31
Eigentlich war es mir nicht wichtig, wie die Geschichte für die Hauptfigur ausgeht.	3.09	1.93	0.70	-0.75
Ich wollte unbedingt, dass die Hauptfigur gut aus der Geschichte herauskommt.	5.05	1.73	-0.72	-0.37
Ich wollte, dass die Hauptfigur unbedingt ihre Ziele erreicht.	4.67	1.73	-0.51	-0.56
Ob die Hauptfigur ihr Ziel erreicht, war mir egal.	3.05	1.84	0.69	-0.64
Mir war egal, wie die Geschichte für die Hauptfigur ausgeht.	2.92	1.92	0.82	-0.53
Für die Hauptfigur sah es lange nicht gut aus.	4.72	1.82	-0.54	-0.73

Das Schicksal der Hauptfigur hing in der Schwebe.	4.74	1.88	-0.59	-0.79
Die Hauptfigur konnte sich lange Zeit nicht sicher sein, ob sie am Ende erfolgreich sein würde.	5.24	1.43	-0.71	-0.32
Der Film hat die Zuschauer lange im Unklaren gelassen, ob alles gut ausgeht.	4.85	1.85	-0.56	-0.81
Der Film ließ lange Zeit offen, wie sich die Geschichte für die Hauptfigur entwickelt.	4.72	1.78	-0.40	-1.00
Die Hauptfigur konnte lange nicht wissen, ob die Geschichte gut für sie ausgeht.	5.30	1.52	-0.99	0.37
Für die Hauptfigur war von Beginn an klar, dass sie ihre Ziele erreichen würde.	3.10	1.90	0.64	-0.76
Die Hauptfigur musste keine Zweifel an der Erreichung ihrer Ziele haben.	2.45	1.38	1.13	0.97
Für die Hauptfigur bestand kein Risiko.	2.35	1.52	1.27	0.98
Für die Hauptfigur sah es stets gut aus.	2.30	1.36	1.16	0.87
Für die Hauptfigur ging es um viel.	5.38	1.69	-0.99	0.04
Für die Hauptfigur stand einiges auf dem Spiel.	5.25	1.71	-0.80	-0.45
Die Hauptfigur hat sehr viel riskiert.	4.71	1.81	-0.35	-1.05
Die Hauptfigur ist ein großes Wagnis eingegangen.	4.45	1.96	-0.26	-1.17
Die Hauptfigur hat alles auf eine Karte gesetzt.	4.44	1.86	-0.23	-0.96
Für die Hauptfigur ging es um wenig.	2.34	1.44	1.33	1.24
Für die Hauptfigur ging es um nichts.	1.99	1.30	1.63	2.34
Für die Hauptfigur stand wenig auf dem Spiel.	2.65	1.72	0.81	-0.47
Die Hauptfigur hatte wenig zu befürchten.	2.42	1.42	1.25	1.39
Die Hauptfigur ging kein besonders großes Risiko ein.	2.75	1.67	0.91	-0.12
Die Hauptfigur hat nichts gewagt.	2.35	1.41	1.14	0.91

Anmerkung. Items der finalen Skala hervorgehoben; $n = 222$.

Ziel der Itemreduktion war einerseits eine konsistente Faktorenlösung und andererseits eine Itemanzahl, die einen sowohl validen und reliablen als auch ökonomischen Einsatz der Skala im Feld ermöglicht und daher zwischen ca. zehn und 15 Items enthalten sollte. Am Ende des geschilderten explorativen Prozesses stand eine Auswahl von zwölf (vier je Dimension) Items.

Tabelle 3. Faktorenlösung mit den zwölf finalen Items

		Faktor			h ²
		1	2	3	
	Präferenz für Handlungsausgang ($\omega = 0.91$)				
P1	Ich habe mir ein gutes Ende für die Hauptfigur gewünscht.	0.91			0.80
P2	Ich wollte unbedingt, dass die Hauptfigur gut aus der Geschichte herauskommt.	0.86			0.77
P3	Ich wollte, dass die Hauptfigur unbedingt ihre Ziele erreicht.	0.87			0.74
P4	Mir war egal, wie die Geschichte für die Hauptfigur ausgeht.	-0.76			0.61
	Unsicherheit ($\omega = 0.82$)				
U1	Die Hauptfigur konnte sich lange Zeit nicht sicher sein, ob sie am Ende erfolgreich sein würde.		0.65		0.61
U2	Der Film hat die Zuschauer lange im Unklaren gelassen, ob alles gut ausgeht.		0.64		0.55
U3	Der Film ließ lange Zeit offen, wie sich die Geschichte für die Hauptfigur entwickelt.		0.85		0.61
U4	Die Hauptfigur konnte lange nicht wissen, ob die Geschichte gut für sie ausgeht.		0.59		0.48
	Risikowahrnehmung ($\omega = 0.84$)				
R1	Für die Hauptfigur ging es um viel.			-0.52	0.40
r2	Die Hauptfigur hat sehr viel riskiert.			-0.67	0.59
R3	Für die Hauptfigur ging es um wenig.			0.82	0.72
R4	Die Hauptfigur ging kein besonders großes Risiko ein.			0.80	0.56
	erklärte Varianz	24,8 %	19,6 %	17,7 %	Σ 62,1 %

Anmerkung. Studie 2, $n = 222$, Hauptachsenanalyse, Faktorenanzahl über Parallelanalyse ermittelt, Oblique Rotation (Promax): $0,36 < r < 0,52$, KMO = 0,85, Faktorladungen $< 0,2$ ausgeblendet, Reliabilität der Gesamtskala: $\omega = 0,89$.

Die Hauptachsenanalyse mit einer nicht orthogonalen Rotation (*Promax*) dieser zwölf Items konnte einerseits die anvisierte Faktorenstruktur spiegeln und lieferte andererseits zufriedenstellende Kennwerte (Tab. 3), sowohl hinsichtlich der Struktur als auch der einzelnen Items.

5.3 Erste Skalenevaluation

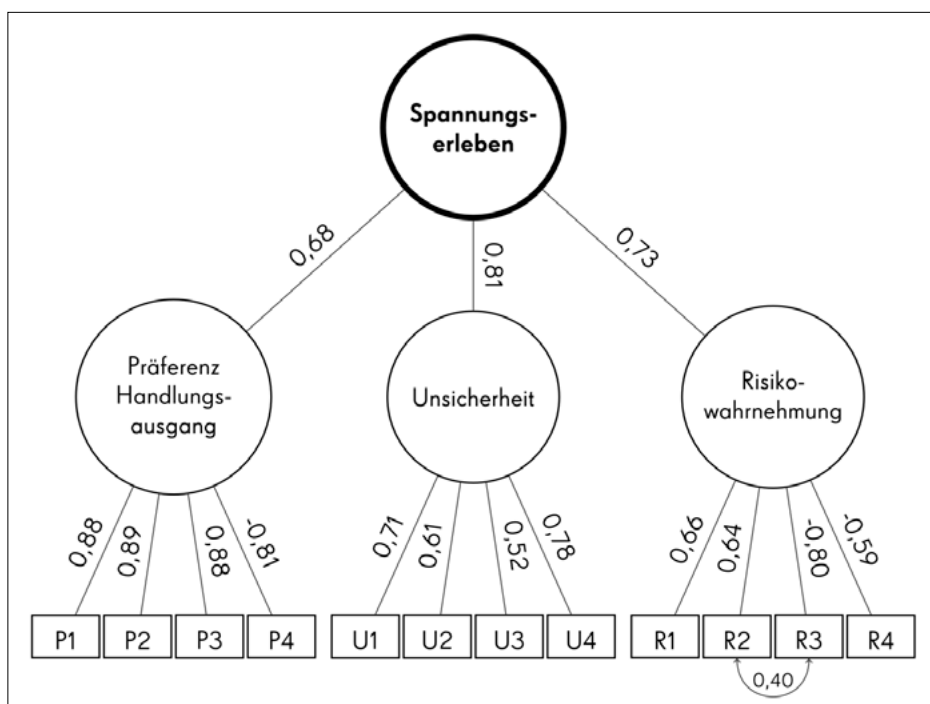
Diese 12-Item-Lösung wurde anschließend an diesen explorativen Prozess der Itemreduktion und Dimensionierung einer (ersten) *konfirmatorischen Faktorenanalyse* (CFA) unterzogen. Die Analyse erfolgte in *jamovi* (The jamovi project, 2022) unter Einsatz des Pakets *SEMLj* in Version 0.9.0 (Gallucci & Jentschke, 2021), basierend auf dem R-Paket *lavaan* (Rosseel, 2019). Als Datenbasis dienten 216 gültige Fälle der hier beschriebenen kombinierten Feld- und Laborstudie. Eingang in das Messmodell fanden die zwölf genannten manifesten Variablen, die als reflektive Indikatoren für die drei latenten Konstrukte dienten – die drei Dimensionen des Konstrukts Spannungserleben. Spannungserleben wiederum wurde als weiteres latentes Konstrukt höherer Ordnung (*second-order factor*; vgl. z. B. Yale et al., 2015) in das Messmodell integriert (die Struktur des Messmodells kann Abb. 1 entnommen werden). Die manifesten Variablen wurden für die Analyse vorab standardisiert; als Schätzmethode kam *Maximum Likelihood* zum Einsatz. Durch Freisetzen eines Parameters ließ sich der anfänglich nur akzeptable Modellfit deutlich verbessern und wurde auf Basis der Modifikationsindizes sowie theoretischer Plausibilität ausgewählt. Dabei handelt es sich um die Kovarianz zweier Indikatoren der Dimension *Unsicherheit* („Der Film hat die Zuschauer lange im Unklaren gelassen, ob alles gut ausgeht.“ und „Der Film ließ lange Zeit offen, wie sich die Geschichte für die Hauptfigur entwickelt.“). Wir vermuten hier einen gemeinsamen (unerwünschten) Effekt der Verwendung des Begriffs „Film“. Es deuten zwar keine sonstigen Befunde auf eine solche Problematik hin³, aber plausibel scheint dennoch eine Irritation befragter Personen, wenn sie den Begriff „Film“ auf eine gerade gesehene Serienepisode anwenden sollen. Zur Modellevaluation werden χ^2 , der Quotient aus χ^2 und *df*, RMSEA, SRMR sowie CFI hinzugezogen (vgl. z. B. Kline, 2015). Die Cut-off-Werte orientieren sich an den von Schermelleh-Engel et al. (2003) zusammengetragenen Vorschlägen. Die verschiedenen Fit-Indizes für das so spezifizierte Modell weisen insgesamt einen angemessenen (RMSEA = 0,055, 95 % CI [0,033; 0,076]) bis guten und sehr guten Fit auf (χ^2 = 82,9; *df* = 50; χ^2/df = 1,66; SRMR = 0,040; CFI = 0,976). Da dieser Test des Messmodells auf den gleichen empirischen Daten beruht, die bereits bei der explorativen Modellspezifikation zum Einsatz kamen, ist die konfirmatorische Aussagekraft nicht ausreichend (Seaman & Weber, 2015). Aus diesem Grund wurde in einem nächsten Schritt diese Analyse an einer weiteren, neu erhobenen und somit unabhängigen Stichprobe wiederholt.

3 Zudem wurden den Befragten auf der ersten Seite des Fragebogens folgende Erläuterung geliefert: „Auf den folgenden Seiten finden Sie Fragen, die sich auf die eben gesehene Serie oder den eben gesehenen Film beziehen. Zur Vereinfachung wird im Folgenden immer von ‚Film‘ gesprochen. Gemeint ist damit aber natürlich auch eine soeben angeschaut Serie.“

6. Studie zwei – konfirmatorische Faktorenanalyse

Die zweite Studie wurde zunächst ausschließlich zum Zweck einer zweiten CFA durchgeführt, daher wurden alle zwölf Items der oben dargestellten Drei-Faktoren-Lösung erhoben. Letztlich fanden aber auch Konstrukte in den Fragebogen Eingang, die der späteren Validierung dienen sollen (bspw. *Sympathie* und *affektive Disposition*; vgl. Abschnitt 7.2). Die Erhebung fand ausschließlich im Labor (im Rahmen einführender kommunikationswissenschaftlicher Lehrveranstaltungen) statt. Insgesamt nahmen an dieser Studie 226 Proband:innen teil, wobei es sich um eine anfallende studentische Stichprobe handelt (Durchschnittsalter 21 Jahre; 75 % weiblich). Als Stimulus kamen neben den bereits in der ersten Laborstudie verwendeten Episoden von drei TV-Serien (*Friends*, $n = 14$; *Simpsons*, $n = 45$; *Grey's Anatomy*, $n = 110$) noch zwei Kurzfilme (*Am Ende der Wald*, $n = 30$; *Das Grüffelo*, $n = 27$) zum Einsatz.

Abbildung 1. Messmodell der drei Dimensionen der Skala Spannungserleben (SSE)



Anmerkung. Es werden standardisierte Parameter ausgewiesen. Alle Faktorladungen sind signifikant ($p < 0,001$). Modell: $\chi^2 = 101$; $df = 50$; $\chi^2/df = 2,02$; RMSEA = 0,068, 95 % CI [0,049; 0,087]; SRMR = 0,047; CFI = 0,967

Die CFA (Abb. 1) folgte exakt dem gleichen Vorgehen und Interpretationsrahmen wie die eben vorgestellte erste CFA. Auch hier erwies sich im Rahmen der Modelltestung wieder eine Optimierung durch die Freisetzung eines Parameters als notwendig und zielführend. Es handelt sich dabei um dieselbe Kovarianz zwischen

zwei Items der Dimension *Unsicherheit*. Dass auch bei dieser unabhängigen, zweiten Stichprobe die empirisch ermittelten Modifikationsindizes auf dieselbe Optimierung verweisen, betrachten wir als einen ersten Hinweis auf die Stabilität des in Studie 1 entwickelten Messmodells. Auch die anderen Befunde zum Modell-Fit bestätigen dies. Wenngleich auch die meisten Kennwerte des Modell-Fits leicht schlechter ausfallen als in Studie 1, verweisen alle mindestens auf einen adäquaten (RMSEA = 0,068, 95 % CI [0,049; 0,087]) oder guten Fit ($\chi^2 = 101$; $df = 50$; $\chi^2/df = 2,02$; SRMR = 0,047; CFI = 0,967).

7. Validierung

Der letzte Schritt in der Entwicklung der Spannungsskala stellt die Validierung dieser dar. Zur Überprüfung der Konstruktvalidität wurden die bereits erläuterten Voraussetzungen (siehe Abschnitt 2.1) und Konsequenzen (siehe Abschnitt 2.3) von Spannung in ein *nomologisches Netzwerk* (Feigl, 1958) eingebettet und die erwarteten Zusammenhänge in Form von Hypothesen ausformuliert.

7.1 Nomologisches Netzwerk & Hypothesen

Geprüft wurde, ob die Spannungsskala positiv mit den Konstrukten *State-* (H1) und *Trait- Empathie* (H2a&b), *Narratives Erleben* (H3), *Identifikation* (H4), *affektive Disposition* (H5) sowie *Sympathie* (H6) korreliert. Darüber hinaus wurde getestet, ob die fünf verschiedenen Stimuli der Laborstudie mit einem unterschiedlichen Spannungserleben einhergehen (H7). Mit der Tendenz, während des Films oder der Serie *umzuschalten* (H8) und *Nebentätigkeiten* nachzugehen (H9), sollte die Skala dagegen negativ korrelieren. Wiederum positive Zusammenhänge wurden für die *Motivation zur Zweitrezeption* der Stimuli (H10) und dessen *Bewertung* (H11) erwartet. Außerdem wurde davon ausgegangen, dass ein positiver Zusammenhang mit dem *Außenkriterium* der subjektiven Bewertung des Spannungsgrades des Stimulus durch die Rezipierenden und der Spannungsskala besteht. Die eindimensionale Messung der erlebten Spannung sollte zu ähnlichen – aber nicht identischen – Ergebnissen wie die SSE kommen (H12).⁴

4 Neben den zwölf erläuterten Validierungskonstrukten erschien zunächst auch die Annahme plausibel, dass das Erleben von Spannung abhängig von der wiederholten Rezeption eines Films oder eine Serie ist. Anders ausgedrückt: Die Proband:innen, die das Ende eines Stimulus nicht kennen, sollten ein höheres Spannungserleben haben, da Spannung per Definition mit der Unsicherheit über den Handlungsausgang einhergeht (siehe Abschnitt 2.2). Die wiederholte Rezeption könnte sich zudem auf das Wahrnehmen von Risiken, welchen die Protagonist:innen im Laufe einer Geschichte ausgesetzt sind, auswirken.

Gleichzeitig wird in der theoretischen Auseinandersetzung mit dem Phänomen Spannung unter dem Begriff *Paradox of Suspense* das intensive Erleben von Spannung *trotz* wiederholter Rezeption diskutiert. Beispielsweise geht Carroll (1996) davon aus, dass hohes Spannungserleben dennoch möglich ist, weil die Zuschauenden während der Rezeption andere Handlungsausgänge simulieren und sich hypothetisch auf diese einlassen, obwohl sie eigentlich wissen, wie die Geschichte ausgeht: „the heroine in all probability is likely to be killed – the recidivist, despite what he or she knows about the last-minute rescue, recognizes a situation in which the good is unlikely“ (Carroll, 1996, S. 87). Aufgrund der widersprüchlichen Argumente rückten wir letztlich von der Verwendung dieses Validierungskonstrukts wieder ab.

7.2 Vorgehen

Die Operationalisierung der zwölf Validierungskonstrukte erfolgte dabei – sofern dies möglich war – auf Basis etablierter Skalen. Für einige Skalen mussten die Items zunächst aus dem Englischen ins Deutsche übersetzt werden. Die (adaptierten) Skalen wurden auf ihre Reliabilität überprüft und erzielten fast durchgehend sehr gute Ergebnisse. Diese wurden als siebenstufige Ratingskalen mit den verbalen Ankern „trifft überhaupt nicht zu“ und „trifft voll und ganz zu“ operationalisiert.

Zur Messung von *State-Empathie*, also dem Erleben situativer Empathie, wurde die Skala von Shen (2010) eingesetzt ($\omega = 0,88$). Empathiefähigkeit wurde durch die *Trait-Empathie* Skala („E-Skala“) von Leibetseder et al. (2001) erfasst ($\omega = 0,92$). Diese begreift Trait-Empathie als zweidimensionales Konstrukt, das Empathie mit Personen in fiktivem ($\omega = 0,93$) oder realem Kontext ($\omega = 0,86$) misst. Die *Narrative Engagement-Skala* von Busselle & Bilandzic (2009) wurde zur Messung von *Narrativem Erleben* eingesetzt ($\omega = 0,81$). Die Skala lag in deutscher Übersetzung vor und fasst das Konstrukt in den vier Dimensionen – *Aufmerksamkeitsfokus*, *narratives Verstehen*, *narrative Präsenz* und *emotionale Beteiligung*. Zur Messung von *Identifikation* wurde Cohens *Identification Scale* (2001) herangezogen. Dabei kam die Hälfte der ursprünglich zehn Items zum Einsatz ($\omega = 0,70$). Bei den Skalen zur Messung von *affektiver Disposition* ($\omega = 0,87$) und *Sympathie* ($\omega = 0,92$) mit jeweils sieben Items handelte es sich um Messinstrumente, die eigens für die Validierung der Spannungsskala entwickelt wurden. Die Operationalisierung der *Werkseigenschaften* der Filme oder Serien erfolgte über den Einsatz verschiedener Stimuli. Die *subjektive Bewertung des Spannungsgrades*, des Bedürfnisses während des Films bzw. der Serie *umzuschalten* oder *Nebentätigkeiten* nachzugehen sowie die Motivation zur *Zweitreception* wurde jeweils als eigenständige Frage mittels einer siebenstufigen Ratingskala operationalisiert, wobei hier die verbalen Verankerungen der Pole an das jeweilige Konstrukt angepasst wurden (z. B. H12: „überhaupt nicht spannend“ und „sehr spannend“). Zur *Bewertung des Films bzw. der Serie* sollten die Proband:innen dem Stimulus eine Schulnote zwischen eins und sechs geben.

Die Validierung der Skala erfolgte mittels der Daten aus Studie 1 und 2 ($n = 448$), wobei einige Hypothesen (H1, H2, H3, H4 und H9) lediglich mit den im Feld erhobenen Daten von Studie 1 ($n = 99$) geprüft wurden. Die Werkseigenschaften der Stimuli (H7) wurden ausschließlich in der Laboruntersuchung innerhalb der Studie 1 ($n = 123$) erhoben. Die Prüfung der Zusammenhänge erfolgte überwiegend auf Basis eines Korrelationstests (Pearson). Hypothese 7 wurde anhand einer ANOVA geprüft.

7.3 Ergebnisse

Insgesamt konnte das nomologische Netzwerk sowie die darauf basierenden Hypothesen – mit der Ausnahme von Hypothese 9 – bestätigt werden. Abbildung 2 gibt einen Überblick über die Ergebnisse der Konstruktvalidierung.

Voraussetzungen von Spannung

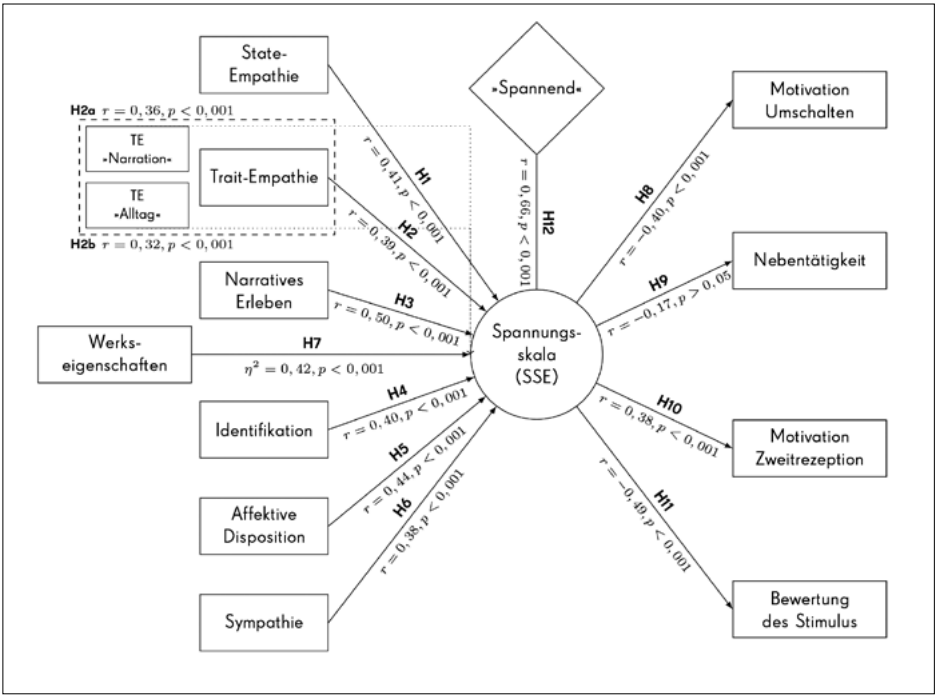
Spannung entsteht, wie in Abschnitt 2 erläutert, wenn die Rezipierenden Unsicherheit in Bezug auf den Handlungsausgang einer Narration empfinden und sie dabei einen bestimmten Handlungsverlauf präferieren. Diese Präferenz ist an die *Empathie* mit den Protagonist:innen gebunden (Carroll, 1996; Zillmann, 1996) und damit „Voraussetzung [...] für das Spannungserleben insgesamt“ (Wünsch, 2014, S. 230). Der Zusammenhang zwischen dem *situativen* Erleben von *Empathie* ($r = 0,41, p < 0,001$) bei der Rezeption bzw. der Fähigkeit zum empathischen Mitfühlen ($r = 0,39, p < 0,001$) und der Spannungsskala konnte wie erwartet bestätigt werden. Auch die beiden (Sub-)Dimensionen von *Trait-Empathie* – die Fähigkeiten zum Erleben von Empathie in sowohl alltäglichen ($r = 0,32, p < 0,001$) als auch narrativen Kontexten ($r = 0,36, p < 0,001$) – weisen einen positiven Zusammenhang mit der SSE auf. Da *Narratives Erleben* für ein tiefes Involvement mit dem Medieninhalt sorgt, lag darüber hinaus die Vermutung nahe, dass das Rezeptionsphänomen ein weiterer Prädiktor für das Erleben von Spannung sein könnte. Die empirischen Daten bestätigen diese Vermutung ($r = 0,50, p < 0,001$). Auch die Annahme, dass Spannung und *identifikatorische Prozesse* – also das intensive Hineinversetzen der Proband:innen in die Charaktere – bei der Medienrezeption miteinander einhergehen, konnte angenommen werden ($r = 0,40, p < 0,001$). Die Präferenz für einen bestimmten Handlungsausgang für die Charaktere einer Narration basiert, wie in Abschnitt 2.1 erläutert, auf der Entwicklung moralischer Wertvorstellung und *affektiven Dispositionen*. Sehen die Zuschauer:innen diese Präferenz in Gefahr, wird der Medieninhalt als spannend erlebt. Dieser Zusammenhang zwischen *affektiven Dispositionen* und der Spannungsskala konnte auf Basis der empirischen Daten auch bestätigt werden ($r = 0,44, p < 0,001$). Ebenso die Annahme, dass *Sympathie* für die Protagonist:innen mit einem erhöhten Spannungserleben einhergeht ($r = 0,38, p < 0,001$). Aus werkstheoretischer Perspektive wurde außerdem vermutet, dass die Merkmale von Medienangeboten (z. B. die technische und dramaturgische Gestaltung, die mit einem spezifischen Stimulus verbunden ist) einen Einfluss auf das Entstehen von Spannung haben. Auch diese Hypothese auf Ebene der Prädiktoren konnte damit angenommen werden ($F(4, 96) = 60,7, p < 0,001, \eta^2 = 0,42$).

Konsequenzen von Spannung

Neben den Prädiktoren von Spannungserleben bildet das nomologische Netzwerk auch Konsequenzen des rezeptiven Spannungserlebens ab. So wurde angenommen, dass die Proband:innen, die viel Spannung erlebten, eine geringe *Motivation* hatten, *umzuschalten* oder *Nebentätigkeiten* nachzugehen. Während Ersteres be-

stätigt werden konnte ($r = -0,40, p < 0,001$), musste die zweite Annahme abgelehnt werden. Hier zeigte sich kein Zusammenhang ($r = -0,17; p > 0,05$). Eine denkbare Erklärung hierfür liegt möglicherweise in den Nebentätigkeiten selbst sowie in unserer Operationalisierung. Bei der Abfrage der Häufigkeit von Nebentätigkeiten wurden den Befragten verschiedene Beispiele solcher Tätigkeiten genannt, von denen aber nicht alle eindeutig einer fehlenden Motivation für den Aufmerksamkeitsfokus auf den Stimulus zuzuordnen sind (z. B.: Handy nutzen, die Toilette besuchen). Daher gehen wir inzwischen von einer nicht vollauf validen Erfassung einer durch fehlende Spannung induzierten Motivation für andere Beschäftigungen aus. Die Widerlegung dieser Hypothese könnte mithin zwar auf eine fehlende Validität unserer Skala hindeuten, für deutlich plausibler halten wir hingegen ein Problem bei unserer Operationalisierung.

Abbildung 2. Nomologisches Netzwerk zur Validierung der Skala Spannungserleben (SSE)



Wie in Abschnitt 2.3 erläutert, wird Spannung zudem als ein Aspekt des *Unterhaltungserlebens* betrachtet (z. B. Czichon, 2019; Lewis et al., 2014; Oliver & Bartsch, 2010; Tamborini et al., 2011; Vorderer & Ritterfeld, 2009; Vorderer et al. 2004) und als Prädiktor von Enjoyment. Entsprechend wurde im Zuge der Validierung angenommen, dass Spannung positiv mit der *Motivation zur erneuten Rezeption des Films oder der Serie* und der *Bewertung des Stimulus* korreliert. Beide Annahmen konnten bestätigt werden ($r = 0,38, p < 0,001$; $r = -0,49, p <$

0,001). Zuletzt konnte auch die Hypothese angenommen werden, dass ein positiver Zusammenhang zwischen der *subjektiven, eindimensionalen Bewertung des Spannungsgrades* und der Skala besteht ($r = 0,66$, $p < 0,001$). Das Messmodell mit nur einem einzigen Indikator kann Spannung also zwar auch erfassen, allerdings ist es nur mit einer Varianzaufklärung von knapp 44 Prozent dazu in der Lage, unser theoretisch und empirisch gesättigtes, multidimensionales Messmodell abzubilden und ist diesem somit deutlich unterlegen.

Insgesamt haben sich elf der zwölf Hypothesen in die vermutete Richtung mit mittelstarken Zusammenhängen bestätigt. Lediglich bei Hypothese 9 zeigte sich kein Zusammenhang mit der Spannungsskala und dem Nachgehen von Nebentätigkeiten, wobei wir die Ursache für die gescheiterte Validierung in einer fehlerhaften Erfassung der hier relevanten Nebentätigkeit sehen. Vor diesem Hintergrund erachten wir unsere Validierung insgesamt als erfolgreich. Damit stellt die *Skala Spannungserlebens (SSE)* ein valides Instrument zur Spannungsmessung dar.

8. Diskussion

Das Ziel der vorliegenden Arbeit war es, basierend auf einer Auseinandersetzung mit dem Forschungsstand, ein Messinstrument zur postrezeptiven Erfassung von Spannung während der Rezeption narrativer Inhalte zu entwickeln.

Im ersten Teil des Beitrags wurde ausgehend von einer umfangreichen Literatursichtung Spannung als dreidimensionales Konstrukt – bestehend aus einer *Präferenz für einen Handlungsausgang*, *Unsicherheit* und *Risikowahrnehmung* – konzeptualisiert. Die theoretische Erörterung wird nunmehr der Mehrdimensionalität des Phänomens (z. B. Borringo, 1980; Carroll, 1990; Schulze, 2006; Zillmann, 1996) gerecht und steht grundsätzlich im Einklang mit einschlägigen Ansätzen der Spannungsforschung (u. a. Borringo, 1980; Carroll, 1990; de Wied, 1991; Eder, 2007; Früh, 2002; Fuchs, 2000; Hant, 1999; Öhding, 1998; Zillmann, 1996). Aufbauend auf die Dimensionalisierung wurde die SSE entwickelt und validiert. Dieser Prozess wurde im Einzelnen dargelegt und umfasste folgende Schritte: Zunächst wurden explorative Vorstudien zur Itemgenerierung durchgeführt. Der Itempool wurde in einer ersten Studie reduziert und evaluiert. Einer zweiten Studie gelang eine Bestätigung des Messmodells im Rahmen einer konfirmatorischen Faktorenanalyse. Den letzten Schritt bildete die Validierung der SSE. Hierfür wurden, um die Konstruktvalidität zu überprüfen, die Voraussetzungen (siehe Abschnitt 2.1) und Konsequenzen (siehe Abschnitt 2.3) von Spannungserleben in ein *nomologisches Netzwerk* (Feigl, 1958) eingebettet und die erwarteten Zusammenhänge in Form von Hypothesen ausformuliert, überprüft und (nahezu) durchgehend hypothesenkonform bestätigt.

Die Entwicklung einer Skala für das Konstrukt *Spannungserleben* war erfolgreich. Mit der *Skala Spannungserleben (SSE)* liegt nun ein valides und reliables Instrument zur Messung von Spannungserleben vor. Die Skala erfasst das *subjektive Erleben* der Rezipient:innen *während* der Rezeption. Die Messung erfolgt postrezeptiv. Ferner misst die Skala Makro-Spannung, also handlungsübergreifende Spannung und nicht Mikro-Spannung, die sich auf einzelne Sequenzen oder Szenen bezieht (u. a. Brewer, 1996, S. 116; Pfister, 2001, S. 147).

Hinsichtlich des empirischen und theoretischen Vorgehens bestehen einige Limitationen: Wir betrachten das Spannungsphänomen aus einer Rezeptionsperspektive und rücken damit ausgewählte kognitive und affektive Prozesse in den Vordergrund. Wie wir in der Herleitung unserer Konzeptspezifikation aber gezeigt haben, stehen diese von uns in den Fokus gerückten Prozesse in einem klar definierten Verhältnis zu zahlreichen verwandten, in anderen theoretischen Beschreibungen von Spannung als essentiell betrachteten Konstrukten. Dies haben wir im Rahmen der Diskussion der Voraussetzung und Konsequenzen für Spannungserleben dargelegt. Somit ist unsere Konstruktdefinition und die darauf aufbauende Operationalisierung von Spannungserleben anknüpfungsfähig an zahlreichen anderen Spannungstheorien.

Aus forschungsökonomischen Gründen konnte die Validierung lediglich an einer begrenzten Bandbreite von Stimuli erfolgen. Des Weiteren ist zu bemerken, dass ebenfalls aus denselben Gründen auf eine anfallende, stark studentische und junge Stichprobe zurückgegriffen wurde. Ob die Stichprobe einen verzerrenden Einfluss mit Blick auf das Sprachverständnis oder die Rezeptionssozialisation hat, wäre zu überprüfen. Letztlich wurde die Skala in verschiedenen Kontexten (Labor und Feld) und für verschiedene Genres getestet, allerdings lediglich im Paper-Pencil-Modus. Eine Validierung der Skala im Online-Modus steht aus.

Wir gehen davon aus, dass sich die Skala bei gleichbleibender Güte im Online-Modus anwenden lässt. Die optische Gestaltung der Skala bliebe ähnlich zum Paper-Pencil-Modus, weshalb hier kaum Unterschiede zu erwarten sind. Eher kann durch den Online-Modus ggf. sogar an Qualität gewonnen oder bei Fehlern (wie doppelte oder fehlende Antworten) interveniert werden. Für den mündlichen Modus sehen wir keine großen Einsatzbereiche. Dennoch ist auch in diesem Modus eine Anwendung der Skala denkbar.

Wir empfehlen, die SSE aufgrund der Validierung an audiovisuellen, narrativen Inhalten nur für eben solche Inhalte einzusetzen. Die Fragen lassen sich in Selbstadministration effizient und flexibel in einer Befragung einsetzen. Auch wenn es sich um insgesamt zwölf Items handelt, raten wir von einer weiteren Kürzung der vorgestellten Skala ab. Die Validität und Reliabilität einer weiter gekürzten Skala wird durch unsere Befunde nicht gestützt. Aus forschungspraktischen Gründen werden dennoch häufig kurze Skalen gesucht. Die Entwicklung einer eigenen Kurzskaala der SSE ist daher künftig wünschenswert. Zukünftig sollte außerdem eine Anpassung der Skala erfolgen, um das Einsatzfeld auf andere Medien(-inhalte) (z. B. Bücher, Hörspiele, Computerspiele) und andere Befragungsmodi und -kontexte zu erweitern. Da für die Entwicklung und Validierung insbesondere abgeschlossene narrative Bewegtbildinhalte eingesetzt wurden, ist die Validierung (und ggf. Anpassung) der Skala für *nicht* abgeschlossene narrative Medieninhalte wünschenswert. Dabei gehen wir davon aus, dass die SSE auch für nicht abgeschlossene Formate wie Fortsetzungsserien adaptiert werden kann.

Das von uns konzeptualisierte Konstrukt Spannungserleben kann zudem grundsätzlich bei der Rezeption beliebiger medialer Inhalte entstehen, solange die drei zentralen Merkmale (Präferenz für einen Handlungsausgang, Unsicherheit und Risikowahrnehmung) gegeben sind. Diese können aber nur während der Rezeption *narrativer* Inhalte auftreten. Spannung kann damit beispielsweise auch

beim Spielen von Videospielen auftreten, solange diese einer Storyline folgen (wie *The Last of Us*). Folglich ließe sich Spannung in diesem Kontext mit einer adaptierten Version unserer Skala messen.

Literatur

- Appel, M., Koch, E., Schreier, M., & Groeben, N. (2002). Aspekte des Leseerlebens: Skalentwicklung [Aspects of the reading experience: Scale development]. *Zeitschrift für Medienpsychologie*, 14(4), 149–154. <https://doi.org/10.1026//1617-6383.14.4.149>
- Bilandzic, H. (2005). Lautes Denken [Method of thinking aloud]. In L. Mikos & C. Wegener (Hrsg.), *Qualitative Medienforschung. Ein Handbuch* (S. 362–370). UVK.
- Borringer, H.-L. (1980). *Spannung in Text und Film: Spannung und Suspense als Textverarbeitungskategorien* [Suspense in text and film: Tention and suspense as text processing categories]. Pädagogischer Verlag Schwann.
- Brewer, W. F. (1996). The nature of narrative suspense and the problem of rereading. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Hrsg.), *Suspense: Conceptualizations, theoretical analyses, and empirical explorations* (S. 107–127). Lawrence Erlbaum Associates.
- Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A model of narrative comprehension and engagement. *Communication Theory*, 18(2), 255–280. <https://doi.org/10.1111/j.1468-2885.2008.00322.x>
- Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, 12(4), 321–347. <https://doi.org/10.1080/15213260903287259>
- Carroll, N. (1990). *The philosophy of horror or paradoxes of the heart*. Routledge. <https://doi.org/10.4324/9780203361894>
- Carroll, N. (1996). The paradox of suspense. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Hrsg.), *Suspense: Conceptualizations, theoretical analyses, and empirical explorations* (S. 71–92). Lawrence Erlbaum Associates.
- Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, 4(3), 245–264. https://doi.org/10.1207/S15327825MCS0403_01
- Czichon, M. (2019). *Kumulierte Serienrezeption. Ein Modell zur Erklärung des Rezeptionssphänomens Binge Watching* [Cumulative series reception. A model to explain the reception phenomenon of binge watching]. Springer Verlag. <https://doi.org/10.1007/978-3-658-26842-8>
- de Wied, M. A. (1991). *The Role of time structures in the experience of film suspense and duration. A study of the effects of anticipation time upon suspense and temporal variations on duration experience and suspense* [Unpublished doctoral dissertation]. University of Amsterdam, Amsterdam, The Netherlands.
- de Wied, M. A., Zillmann, D., & Ordman, V. (1994). The role of empathic distress in the enjoyment of cinematic tragedy. *Poetics*, 23(1-2), 91–106. [https://doi.org/10.1016/0304-422X\(94\)00010-4](https://doi.org/10.1016/0304-422X(94)00010-4)
- Eder, J. (2007). *Dramaturgie des populären Films: Drehbuchpraxis und Filmtheorie* [Dramaturgy of popular film: screenwriting practice and film theory]. LIT Verlag. <https://doi.org/10.25969/mediarep/759>
- Feigl, H. (1958). The “mental” and the “physical”. In H. Feigl, M. Scriven, & G. Maxwell (Hrsg.), *Minnesota studies in the philosophy of science* (S. 370–497). University of Minnesota Press.

- Früh, W. (2002). *Unterhaltung durch das Fernsehen: Eine molare Theorie* [Entertainment through television: A Molar Theory]. UVK.
- Früh, W., & Frey, F. (2014). Das Narrative und seine Attraktivitäts- und Wirkungspotenziale. Wirkungen der Definitionskomponenten ‚Zielgerichtetheit‘ und ‚menschlicher Handlungsträger‘ bei der Filmrezeption [Narrative and its potential for attraction and effect: Effects of the definitional components ‘purposefulness’ and ‘human agent’ in film reception]. In W. Früh & F. Frey (Hrsg.), *Narration und Storytelling: Theorie und empirische Befunde* (S. 319–371). Halem.
- Fuchs, A. (2000). *Dramatische Spannung: Moderner Begriff – antikes Konzept* [Dramatic suspense: Modern term – ancient concept]. J.B. Metzler. <https://doi.org/10.1007/978-3-476-02751-1>
- Gallucci, M., & Jentschke, S. (2021). SEMlj: jamovi SEM Analysis. [jamovi module]. <https://semlj.github.io/>
- Giles, C. D. (2002). Parasocial interaction: A review of the literature and a model for future research. *Media Psychology*, 4(3), 279–305. https://doi.org/10.1207/S1532785XMEP0403_04
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Green, M. C., & Brock, T. C. (2002). In the mind’s eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, & T. C. Brock (Hrsg.), *Narrative impact: Social and cognitive foundations* (S. 315–341). Lawrence Erlbaum Associates.
- Green, M. C., Brock, T. C., & Kaufmann, G. F. (2004). Understanding media enjoyment: The role of transportation into narrative worlds. *Communication Theory*, 14(4), 311–327. <https://doi.org/10.1111/j.1468-2885.2004.tb00317.x>
- Hant, C. P. (1999). *Das Drehbuch. Praktische Filmdramaturgie* [The screenplay: Practical film dramaturgy]. Hübner Verlag.
- Hartmann, T., Schramm, H., & Klimmt, C. (2004). Personenorientierte Medienrezeption: Ein Zwei-Ebenen-Modell parasozialer Interaktionen [Person-oriented media reception: A two-level model of parasocial interactions]. *Publizistik*, 49(1), 25–47. <https://doi.org/10.1007/s11616-004-0003-6>
- Hastall, M. R. (2014). Spannung [Suspense]. In C. Wünsch, H. Schramm, V. Gehrau, & H. Bilandzic (Hrsg.), *Handbuch Medienrezeption* (257–272). Nomos. https://doi.org/10.5771/9783845260389_257
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach’s alpha for estimating reliability. But.... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Hoffner, C. A., & Levine, K. J. (2005). Enjoyment of mediated fright and violence: A meta-analysis. *Media Psychology*, 7(3), 207–237. https://doi.org/10.1207/S1532785XMEP0702_5
- Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observation on intimacy at a distance. *Journal of Psychiatry*, 19(3), 215–229. <https://doi.org/10.1080/00332747.1956.11023049>
- Kintsch, W. (1993). Information accretion and reduction in text processing: Inferences. *Discourse Processes*, 16(1-2), 193–202. <https://doi.org/10.1080/01638539309544837>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford Press.
- Knobloch, S., Patzig, G., Mende, A.-M., & Hastall, M. (2004). Affective news: Effects of discourse structure in narratives on suspense, curiosity, and enjoyment while reading news and novels. *Communication Research*, 31(3), 259–287. <https://doi.org/10.1177/0093650203261517>

- Knobloch-Westerwick, S., David, P., Eastin, M. S., Tamborini, R., & Greenwood, D. (2009). Sports spectators' suspense: Affect and uncertainty in sports entertainment. *Journal of Communication*, 59(4), 750–767. <https://doi.org/10.1111/j.1460-2466.2009.01456.x>
- Leibetseder, M., Laireiter, A.-R., Riepler, A., & Köller, T. (2001). E-Skala: Fragebogen zur Erfassung von Empathie – Beschreibung und psychometrische Eigenschaften [E-scale: Questionnaire for the assessment of empathy – description and psychometric properties]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 22(1), 70–85. <https://doi.org/10.1024//0170-1789.22.1.70>
- Lewis, R. J., Tamborini, R., & Weber, R. (2014). Testing a dual-process model of media enjoyment and appreciation. *Journal of Communication*, 64(3), 397–416. <https://doi.org/10.1111/jcom.12101>
- Mahrt, M. (2014): Einstellung [Attitude]. In: C. Wünsch, H. Schramm, V. Gehrau, & H. Bilandzic (Hrsg.), *Handbuch Medienrezeption* (S. 113–127). Nomos. https://doi.org/10.5771/9783845260389_113
- Mikos, L. (2008). Trauer, Wut und Leidenschaft. Empathie bei Film und Fernsehen [Grief, anger and passion. Empathy in film and television]. *TV-Diskurs*, 12(1), 18–23.
- Möhring, W., & Schlütz, D. (2013). Standardisierte Befragung. Grundprinzipien, Einsatz und Anwendung [Standardized survey: Basic principles, use and application]. In W. Möhring & D. Schlütz (Hrsg.), *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft* (S. 183–200). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-18776-1_10
- Oliver, M. B., & Bartsch, A. (2010). Appreciation as audience response: Exploring entertainment gratifications beyond hedonism. *Human Communication Research*, 36(1), 53–81. <https://doi.org/10.1111/j.1468-2958.2009.01368.x>
- Öhding, B.-K. (1998). *Thriller der neunziger Jahre: Über den Zusammenhang von Struktur, Spannung und Bedeutung an ausgewählten Spielfilmen* [Thrillers of the nineties: On the relationship between structure, suspense and meaning in selected feature films]. Wissenschaftler-Verlag.
- Pfister, M. (2001). *Das Drama: Theorie und Analyse* [The drama: Theory and analysis]. UTB. <https://doi.org/10.36198/9783825205805>
- Rosseel, Y. (2019). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schneider, F. M., Bartsch, A., & Oliver, M. B. (2019). Factorial validity and measurement invariance of the appreciation, fun, and suspense scales across US-American and German samples. *Journal of Media Psychology: Theories, Methods, and Applications*, 31(3), 149–156. <https://doi.org/10.1027/1864-1105/a000236>
- Schulze, A.-K. (2006). *Spannung in Film und Fernsehen: Das Erleben im Verlauf* [Suspense in film and television: The experience during the process]. Logos-Verlag.
- Schwan, S. (2014). Verstehen [Comprehension]. In C. Wünsch, H. Schramm, V. Gehrau, & H. Bilandzic (Hrsg.), *Handbuch Medienrezeption* (S. 191–206). Nomos. https://doi.org/10.5771/9783845260389_191
- Seaman, C. S., & Weber, R. (2015). Undisclosed flexibility in computing and reporting structural equation models in communication science. *Communication Methods and Measures*, 9(4), 208–232. <https://doi.org/10.1080/19312458.2015.1096329>
- Segal, E. M. (1995). Narrative comprehension and the role of deictic shift theory. In J. F. Duchan, G. A. Bruder, & L. E. Hewitt (Hrsg.), *Deixis in narrative: A cognitive science perspective* (S. 3–17). Psychology press. <https://doi.org/10.4324/9780203052907-7>

- Shen, L. (2010). On a scale of state empathy during message processing. *Western Journal of Communication*, 74(5), 504–524. <https://doi.org/10.1080/10570314.2010.512278>
- Tamborini, R., Grizzard, M., David Bowman, N., Reinecke, L., Lewis, R. J., & Eden, A. (2011). Media enjoyment as need satisfaction: The contribution of hedonic and nonhedonic needs. *Journal of Communication*, 61(6), 1025–1042. <https://doi.org/10.1111/j.1460-2466.2011.01593.x>
- The jamovi project (2022). jamovi (Version 2.3) [Computer Software]. <https://www.jamovi.org>
- Thissen, B. A. K., Menninghaus, W., & Schlotz, W. (2021). The pleasures of reading fiction explained by flow, presence, identification, suspense, and cognitive involvement. *Psychology of Aesthetics, Creativity, and the Arts*, 15(4), 710–724. <https://doi.org/10.1037/aca0000367>
- Vorderer, P. (1994). „Spannung ist, wenn’s spannend ist“: Zum Stand der (psychologischen) Spannungsforschung [“Suspense is when it’s suspenseful”: On the state of (psychological) suspense research]. *Rundfunk und Fernsehen*, 42(3), 323–339.
- Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory*, 14(4), 388–408. <https://doi.org/10.1111/j.1468-2885.2004.tb00321.x>
- Vorderer, P., & Ritterfeld, U. (2009). Digital games. In R. L. Nabi & M. B. Oliver (Hrsg.), *The SAGE handbook of media processes and effects* (S. 455–467). Sage Publications.
- Weinacht, S., & Hohlfeld, R. (2007). Das Hofnarren-Komplott: Deskriptiv-theoretische Herleitung von Entgrenzung und Selbstthematization im Journalismus [The court jester scheme: Descriptive-theoretical derivation of dissolution of boundaries and self-thematization in journalism]. In A. Scholl, R. Renger, & B. Blöbaum (Hrsg.), *Journalismus und Unterhaltung: Theoretische Ansätze und empirische Befunde* (S. 157–207). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90724-6_9
- Wünsch, C. (2014). Empathie und Identifikation [Empathy and identification]. In C. Wünsch, H. Schramm, V. Gehrau, & H. Bilandzic (Hrsg.), *Handbuch Medienrezeption* (S. 223–241). Nomos. https://doi.org/10.5771/9783845260389_223
- Wulff, H. J. (1993). Spannungsanalyse: Thesen zu einem Forschungsfeld [Suspense analysis: Theses on a field of research]. In: *montage/av*, 2(2), 97–100.
- Yale, R. N., Jensen, J. D., Carcioppolo, N., Sun, Y., & Liu, M. (2015). Examining first- and second-order factor structures for news credibility. *Communication Methods and Measures*, 9(3), 152–169. <https://doi.org/10/ggxxg45>
- Zillmann, D. (1980). Anatomy of suspense. In P. H. Tannenbaum (Hrsg.), *The entertainment functions of television* (S. 133–163). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781315803050>
- Zillmann, D. (1991). Empathy: Affect from bearing witness to the emotions of others. In J. Bryant & D. Zillmann (Hrsg.), *Responding to the screen: Reception and reaction processes* (S. 135–167). Lawrence Erlbaum Associates, Inc. <https://doi.org/10.4324/9780203812181>
- Zillmann, D. (1996). The psychology of suspense in dramatic exposition. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Hrsg.), *Suspense: Conceptualizations, theoretical analyses, and empirical explorations* (S. 199–231). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203811252>
- Zillmann, D. (2004). Emotionspsychologische Grundlagen [Fundamentals of emotion psychology]. In R. Mangold, P. Vorderer, & G. Bente (Hrsg.), *Lehrbuch der Medienpsychologie* (S. 101–128). Hogrefe.

Anhang

Die Spannungsskala (SSE)

Instruktion:

Auf den folgenden Seiten finden Sie Fragen, die sich auf die eben gesehene Serie oder den eben gesehenen Film beziehen. Zur Vereinfachung wird im Folgenden immer von „Film“ gesprochen. Gemeint ist damit aber natürlich auch eine soeben angeschaute Serie.

[...]

In einigen der folgenden Fragen wird es um die „Hauptfigur“ des Films gehen. Wir meinen damit die für die Handlung wichtigste, Ihnen aber auch sympathische Figur – sozusagen die Filmheldin oder den Filmhelden. Sollte es in dem eben gesehenen Film mehrere Figuren geben, auf die das zutrifft, denken Sie bitte immer an die Ihnen sympathischste. Können Sie uns bitte als Erstes aufschreiben, wer in dem Film eben die Hauptfigur war? Falls Sie den Namen der Figur nicht mehr wissen, reicht auch eine kurze Umschreibung der Figur aus.

[offene Antwortmöglichkeit]

[...]

Im Folgenden interessiert uns, wie Sie sich während des Films gefühlt haben, wie Sie die Geschichte des Films empfunden haben und was Ihnen während des Sehens durch den Kopf ging. Wir haben dazu eine ganze Reihe von Aussagen formuliert, mit denen man das beschreiben kann. Geben Sie bitte immer an, ob die Aussage voll und ganz zutrifft oder überhaupt nicht zutrifft. Mit den Werten dazwischen können Sie Ihr Urteil abstufen.

Items:⁵

P1: Ich habe mir ein gutes Ende für die Hauptfigur gewünscht.

P2: Ich wollte unbedingt, dass die Hauptfigur gut aus der Geschichte herauskommt.

P3: Ich wollte, dass die Hauptfigur unbedingt ihre Ziele erreicht.

P4: Mir war egal, wie die Geschichte für die Hauptfigur ausgeht. (-)

U1: Die Hauptfigur konnte sich lange Zeit nicht sicher sein, ob sie am Ende erfolgreich sein würde.

U2: Der Film hat die Zuschauer lange im Unklaren gelassen, ob alles gut ausgeht.

U3: Der Film ließ lange Zeit offen, wie sich die Geschichte für die Hauptfigur entwickelt.

U4: Die Hauptfigur konnte lange nicht wissen, ob die Geschichte gut für sie ausgeht.

R1: Für die Hauptfigur ging es um viel.

R2: Die Hauptfigur hat sehr viel riskiert.

R3: Für die Hauptfigur ging es um wenig. (-)

R4: Die Hauptfigur ging kein besonders großes Risiko ein. (-)

- 5 Die Skalierung der Items wurde wie folgt angelegt: 7-stufige diskrete Ratingskala mit visueller sowie verbaler Verankerung mit den Extrema „1 = trifft überhaupt nicht zu“ und „7 = trifft voll und ganz zu“ (s. Abschnitt 4 und 5).

Auswertung:

Wir empfehlen die einzelnen Itemwerte als Mittelwertindex auszuwerten. Es können einerseits Mittelwertindizes als Subwerte für die drei Dimensionen (Präferenz für einen Handlungsausgang, Unsicherheit und Risikowahrnehmung) ausgewiesen werden. Andererseits kann, aufgrund der Reliabilität der Skala insgesamt, ein Mittelwertindex aus den drei Indizes der einzelnen Dimensionen für die gesamte Skala Spannungserleben ausgewiesen werden.