

Stage 2 Registered Report

Exploring the Consistency and Value of Humour Style Profiles

Abstract

Establishing generalisable humour style profiles promises to have significant value for educational, clinical, and occupational application. However, previous research investigating such profiles has thus far presented inconsistent results. To determine the generalisability and value of humour style profiles, a large and geographically diverse examination of humour styles was conducted through a cross-sectional questionnaire methodology involving 863 participants from across three world regions. Findings identify inconsistencies in the humour style profiles across countries tested and the extant literature, possibly indicative of cultural differences in the behavioural expression of trait humour. Furthermore, when directly compared, humour types, rather than humour styles, consistently provide the greatest predictive value for friendship and well-being outcomes. As such, with respect to both consistency and value, capturing humour style profiles appears to represent a relatively reductionist approach to appreciating the nuances in the use and consequences of humour.

Key words: Humour; Humour Styles; Cluster Analysis; Culture

Introduction

Humour is “a verbal or non-verbal social communicative event which is purposely initiated to amuse an ‘audience’, or which unintentionally becomes perceived as amusing” (Evans & Steptoe-Warren, 2018, p. 443). Humour has been noted as an important communicative device (Huang & Kuo, 2011; Li & Seale, 2007; Schnurr & Chan, 2009; Wanzer, Wojtaszczyk, & Kelly, 2009). Whilst often considered ideologically positive (Billig, 2005, p. 10), humour has been associated with a diverse range of positive and negative outcomes. As such, humour has been meta-analytically linked to variation in mental health (Schneider, Voracek, & Tran, 2018), work (Mesmer-Magnus, Glew, & Viswesvaran, 2012), advertising (Eisend, 2009), and relationship (Hall, 2017) outcomes.

The most noteworthy development in the field of humour research has been the classification of different types of humour. The most popular of which, is that of Martin et al. (2003). Conceptualising two key distinctions, the target (self/relationships) and valence of the humour (benign/negative), four humour types were proposed: affiliative, aggressive, self-enhancing, and self-defeating (Martin et al., 2003). Linked to various psychological outcomes in research fields spanning the field of psychology, the classification of these humour types has received substantial validation (e.g., Martin & Ford, 2018; McCosker & Moran, 2012).

Affiliative humour is the prototypical humour type, representing the use of benign (non-hostile and tolerant) humour to enhance relationships with others (Martin et al., 2003). Affiliative humour has been linked to greater friendship initiation and social competence (Yip & Martin, 2006), self-esteem (Stieger, Formann, & Burger, 2011; Yue, Liu, Jiang, & Hiranandani, 2014) and communication and creativity at work (Evans & Steptoe-Warren, 2018).

Self-defeating humour is less benign and targets the self, as it is used to enhance relationships with others at the expense of the self (Martin et al., 2003). Outcomes associated include increased depressive symptoms (Tucker et al., 2013) and lower self-esteem (Leist & Müller, 2013) and intimacy (Kuiper, Kirsh, & Maiolino, 2016).

Aggressive humour represents use of less benign humour targeted at relationships in order to enhance oneself often at the expense of others (Martin et al., 2003). Unsurprisingly, aggressive humour use has been associated with greater antagonism, disinhibition and aggression (Martin et al., 2003; Zeigler-Hill, McCabe, & Vrabel, 2016), and lower happiness (Ford, McCreight, & Richardson, 2014).

Self-enhancing humour is the benign humour used to enhance ones’ self, and it is often considered a type of coping or emotion regulation (Hughes & Evans 2016; Kuiper, Martin, & Olinger, 1993; Martin, 1996). As such, self-enhancing humour is often negatively associated with anxiety (Ford, Lappi, O’Connor, & Banos, 2017) and depressive symptoms (Tucker et al., 2013) and positively associated with psychological well-being factors like self-esteem and life satisfaction (Leist & Müller, 2013).

A large amount of cross-sectional work has shown that these different humour types correlate to a different extent with key psychological outcomes (see Martin & Ford, 2018, for a thematic overview). Indeed, differentiating between humour types has resulted in significant developments in understanding of humour and how it is applied in occupational, clinical, and educational fields.

However, there has been growing concern over the quality and complexity of analyses conducted in the field (Robert & Yan, 2007). Leist and Müller (2013) specifically raised concerns that considering these four humour types as distinct might not do justice to the complexity of the phenomena. Instead, they proposed that individuals use all four humour types in different ways, and that their combination of use may determine outcomes that are distinct from those associated with use of any individual humour type. Characteristic profiles of use of the four humour types, referred to as humour style, have been argued to present an important way of advancing understanding of humour (Galloway, 2010). Previous research (see Table 1) has proposed several of these humour style profiles using cluster analysis to create groups of individuals who employ the four humour types in a similar manner.

The literature on humour style profiles so far has produced mixed insights. On the one hand, some extent of consistency in humour style profiles can be observed. In particular, there are two humour styles that have emerged across different studies: a) individuals who use all types of humour more than the sample average (typically referred to as humour endorsers); and b) those that use benign humour types more than the average and negative types less than the average (often termed self-enhancers). Differences in outcomes associated with these two groups are detectable, theoretically-relevant, and thus appear meaningful. For example, concurrent with the general trend of the literature highlighted, all papers adopting cluster analysis indicated significant differences in psychological health between these humour styles (see Table 1). In particular, the 'self-enhancer' profile is often associated with the most optimal psychological well-being outcomes in comparison to other humour style profiles. Together these findings indicate that the study of humour style profiles represents a promising avenue of exploration.

On the other hand, there are a number of findings which cast doubt on the consistency and thus value of humour style profiles. For example, the number and profile of the styles identified varies substantively (see Table 1). Acknowledging the two consistent profiles across all analyses, only two papers identify exactly the same profiles of humour style. Whilst this could be considered indicative of a problematic field, the differences identified are not inherently contradictory. There are a variety of possible reasons why humour style profiles may have varied. Four key factors seem likely contributors to such findings.

First, the researcher's decisions surrounding analysis and interpretation could impact the number and profile of styles discussed. For example, Evans and Steptoe-Warren (2018) examined both three- and four-profile solutions; however, due to parsimony, interpretability, and similarities to previous solutions, they only chose to analyse in detail the former. Different practices surrounding whether and how alternative profile solutions are calculated and discussed may have contributed to divergent results.

Second, there are diverse recommendations for sample size for cluster analysis (e.g., Dolnicar, Grun, Leisch, & Schmidt, 2013; Formann, 1984) and thus sample size has dramatically varied from $n = 202$ to $n = 1252$. Sample size can influence cluster formation, with smaller samples typically demonstrating greater within-group variance and less between-group variance, therefore representing less well-defined groups. As such, variation in sample size is likely to have impacted the number, and possibly the profile, of styles.

Third, there are significant cultural differences in humour production and appreciation (Chen & Martin, 2007; Martin & Sullivan, 2013; Yue, Jiang, Lu, & Hiranandani, 2016). For example, Chinese individuals report using significantly less aggressive humour than those from Canada (Chen & Martin, 2007). The extent to which culture may impact humour style profiles is yet unknown.

Finally, the age of samples explored vary from children to adults. As personality, and thus trait humour use, varies across age (Bariaud, 1989; Greengross, 2013; Roberts, Walton, & Viechtbauer, 2006), it is possible that some differences in humour style profiles, particularly those between child and adult samples, could be attributable to developmental differences. In sum, the differences in method and analysis, size, culture, and age of sample are all possible contributors to inconsistencies in the findings reported so far. This poses the question as to whether there is a consistent underlying taxonomy of humour styles.

The first focus of the current study is to address all four of these factors to determine the consistency of humour style profiles. First, this study follows a preregistered analysis plan in accordance with recent recommendations of best practices in science to limit researcher degrees of freedom in the analysis (Munafò, et al., 2017; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). That is, by preregistering all analysis steps, this study addresses possible flexibility that might have driven the inconsistencies across previous works. Second, this study will test a sample size that exceeds best-practice recommendations for cluster analysis (Dolnicar et al., 2013), thereby minimising the influence of high variation regarding the number and profile of styles that is typically the case with small samples. Third, this study will recruit samples from three countries to determine whether previous inconsistencies in humour style profiles are likely attributable to cultural influences. Fourth, this study only recruits adults, to control for the difference in development of humour styles between adolescents and adults.

In addition to testing the consistency of humour style profiles by addressing these limitations, the current study has a second focus. Namely, we aim to test to what extent humour style can provide predictive value over and above individual humour types. So far, only Leist and Müller (2013) examined their comparative value. The authors found style group membership to be a stronger predictor of psychological well-being than individual humour types in all analyses (excluding the prediction of flexible goal adjustment by self-enhancing humour). However, their study dichotomised scores on individual humour types, which is problematic as the dichotomisation of ordinal data can often lead to distorted or misleading results (MacCallum, Zhang, Preacher, & Rucker, 2002). Therefore, the predictive power of humour style profiles compared to individual humour types is unclear. The current study thus aims to provide a clearer picture of this predictive power by treating individual humour types as continuous in the analysis. We assess the value of humour types compared to humour style profiles by predicting friendship quality and three well-established indicators of well-being: psychological health, self-esteem, and life satisfaction.

Should humour style profiles demonstrate consistency and add incremental validity over humour types, the taxonomy could be of significant value in clinical, occupational and educational fields following further replication. For example, gaining an understanding of an individual's humour profile may facilitate more appropriate recommendations for coping within counselling. Similarly, taxonomies could become the basis of a more individualised approach to humour interventions or could provide the structure for educational material about individual differences in communication strategies. Should humour styles provide no incremental validity over individual humour types, they likely represent a reductionist understanding of trait humour and social interaction.

Table 1: Similarities and Differences in Extant Humour Style Profiles

Reference	N	Rated	Population	Above average AF, AG, SE and SD (Humour Endorser)	Below average AF, AG, SE and SD (Humour Denier)	Above average AF and SE, below average AG and SD	Above average AG and SD and below average AF and SE	Other profiles	Outcomes examined
Galloway, 2010	318	Self	Australian, adults	1	2	3	4		Personality**, Self-esteem**
Leist & Müller, 2013	305	Self	German, adults	1	2	3			Self-esteem*, Tenacious goal pursuit*, Flexible goal adjustment*, Life satisfaction*
Chang et al., 2015	1252	Self	Chinese, adolescents	1	2	3	4		Demographics*, Stress* Academic Performance*
Fox, Hunter & Jones, 2016	1108	Self	United Kingdom, children	3		4		1 (Above average AG and AF, below average SD and SE) 2. (Above average SD and below average AF, SE and AG)	Depressive Symptoms*, Loneliness*, Self-esteem*
Sirigatti et al., 2016	244	Self	Italian, adults	3		2		1 (Average SD, below average AF, AG and SE)	Psychological well-being*
Evans & Steptoe-Warren, 2018	202	Manager	United Kingdom, adults	3		1		2 (Above average AG and below average AF, SE and SD)	Job satisfaction*, Communication*, Stress*, Creativity*, Leader power*

Note: AF = Affiliative Humour; AG = Aggressive Humour; SE = Self-Enhancing Humour; SD = Self-Defeating Humour;

* = significant differences between styles; ** = differences between styles not tested directly

Aims

To determine the generalisability and value of humour style profiles, a large and geographically diverse examination of humour styles was planned. The scale and scope of the study supports development of the most comprehensive picture of the styles of humour use, acknowledging the role of culture and challenging current inconsistencies in humour style profiles. Furthermore, based upon the focus of previous works, the current research explores the value of such styles for friendship quality and psychological well-being in the form of psychological health, self-esteem, and satisfaction with life.

Method**Procedure**

A cross-sectional correlational design was implemented. Participants were asked to complete demographic questions (age, sex, country of origin, country of residence, education level) and the battery of proposed questionnaires online. Each individual laboratory obtained ethical approval to conduct the study from their IRB unless their institution did not require approval, or the work could be covered by pre-existing approval.

Materials

Affiliative, Aggressive, Self-enhancing and Self-defeating humour was assessed through the Humour Styles Questionnaire (Martin et al., 2003). Each scale has eight items responded to using a seven-point likert ranging from “totally agree” to “totally disagree”. Internal reliability for the scales vary between .77 and .81 (Martin et al., 2003) and the scale has demonstrated consistent factor structures across cultures (e.g., Chen & Martin, 2007).

The 5-item World Health Organization Well-being Index (WHO-5; World Health Organisation, 1998) was adopted to capture health-related subjective well-being. Participants respond to items on a 6-point likert, ranging from ‘All of the time’ to ‘At no time’. The scale has been translated into over 30 languages, is unidimensional with an internal reliability often reported above .9 (e.g., Hajos et al., 2013), and each item adds unique information regarding the level of well-being (Blom, Bech, Högberg, Larsson, & Serlachius, 2012). Topp et al. (2015) conducted a systematic review of the WHO-5, concluding it to be a simple and sensitive measure of well-being, evidencing key practical utility through its use as an outcome measure in clinical trials, and predictive validity with depression.

The Rosenberg Self-Esteem Scale (Rosenberg, 1965) is a 10-item questionnaire to assess global self-worth, scored on a 4-point likert ranging ‘Strongly Agree’ to ‘Strongly Disagree’. The scale has been widely used cross-culturally (Schmitt & Allik, 2005).

The Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985) is a 5-item questionnaire scored on a 7-point likert ranging from ‘strongly agree’ to ‘strongly disagree’. The scale is unidimensional (Atienza, Balaguer, & García-Merita, 2003) is mostly comparable across cultures (Whisman & Judd, 2016) and considered a ‘gold-standard’ (Kaczmarek, Bujacz, & Eid, 2015) in life-satisfaction measurement due to its psychometric qualities and predictive validity e.g. for suicide (Pavot & Diener, 2008).

To acknowledge the social nature of humour, friendship quality was measured by the Inventory of Parent and Peer Attachment (Armsden & Greenberg, 1987). The 10-item peer trust scale was adopted (Wilkinson & Goh, 2014), rated on a 5-point likert. This seemed appropriate given that it would tap into an underlying 'secure attachment' factor (Wilkinson & Goh, 2014). Internal reliability of the original scales has varied from .48 to .96 with 3-week test-retest reliability ranging from .86 to .93 (Wilson & Wilkinson, 2012). It is the most commonly used peer/parent attachment measure (Gorrese & Ruggieri, 2012).

Translation

Where possible, existing validated translated versions of the aforementioned materials were adopted. As Dutch and Polish versions of the friendship quality scale were not available, best practice guidelines for back-translation were adopted (Brislin, 1970). Two bilingual translators translated materials from English to the target language, and two further translators translated this back to English. The translators and study lead discussed and resolved discrepancies before being tested upon two non-academic individuals fluent in the target language. These external readings noted no further misunderstandings and did not require further external reading. Following accepted translations, data collection labs were asked to identify any relevant adjustments necessary for their specific participant sample to be approved by the study lead. No such changes were made.

Sample Size

Cluster Analysis will always create groups regardless of sample size and thus determining preferable size is typically problematic (Sarstedt & Mooi, 2014). Rules of thumb for sample size have ranged from 2 per variable (Formann, 1984) to 70 (Dolnicar et al., 2013), the latter of which has only been exceeded by two previous studies in this field (Galloway, 2010; Leist & Müller, 2013). Because there are four humour types and 70 participants per cluster is the most conservative recommendation, a sample of 280 responses per world region was targeted. To account for possible exclusions (see below), we set the target sample size at exactly 300 completed surveys per world region. There are substantial differences between regions in the number of authors represented and thus subsequent capacity for data collection. As such, an increased target of 500 was set for the United Kingdom to maximise the data available for secondary analyses. To minimise researcher degrees of freedom, the following stopping rules were applied: All authors specified a start date of data collection and an end date that fell five months later. Data collection stopped when either the maximum sample size of 300/500 participants was reached, or at the end date if participant recruitment had been problematic. Start and end date were registered on the Open Science Framework page of this project, and all targets were met before the end date.

Participants

Data was collected from three areas: The United Kingdom (UK), Netherlands (NL), and Poland (PL), representing a total possible sample size of 1100. As the measurement of humour styles reflects a trait-like examination of humour, only individuals over 18 were recruited to develop an adult taxonomy as traits undergo significant development in adolescence (Bariaud, 1989; Greengross, 2013; Roberts, Walton, & Viechtbauer, 2006).

Exclusion

First, participants who did not finish the survey were considered as withdrawn and thus excluded. All questions forced a response to ensure no missing data. This led to an initial sample of 465 for the UK, 300 for the Netherlands, and 234 for Poland.

Second, we followed recent recommendations by Leiner (2013) to obtain high-quality survey data by excluding respondents with a Relative Speed Index (RSI) of > 1.75 . The RSI identifies cases of meaningless data based on the time it takes participants to complete the survey. However, rather than being based on the absolute completion time, the RSI takes into account that absolute completion time can a) easily be skewed by outliers and b) not be compared across studies, as it depends on the length and complexity of the survey. The RSI was computed by dividing the sample's median completion time for each page by the individual participant's page completion time. The resulting factor is known as a speed factor and indicates how fast or slow a respondent went through a specific page in relation to the entire sample. Afterwards, these speed factors were trimmed to an interval of $[0|3]$, which serves to help the researcher to not exclude participants who accidentally skipped a page, but otherwise produced valid data. Last, the trimmed speed factors were averaged to create the RSI. This process excluded 31 of the UK, 8 of the Dutch, and 6 of the Polish participants.

Third, we excluded participants who failed an attention check. The attention check represented one item in the humour scale that read: "To make sure you are paying attention, please select 'Somewhat Disagree'". 24 of the UK, 14 of the Dutch and 45 of the Netherlands sample were excluded on this basis.

Sample Demographics

The final sample included 863 participants.

The UK sample ($n = 410$) had a mean age of 25.8 ($SD = 10.7$), ranging from 18 to 72, and the majority were educated up to A-levels ($n = 228$, 56%) or Undergraduate study ($n = 76$, 19%). The majority ($n = 322$; 79%) were female, with 86 males (21%) and 2 participants not disclosing sex. Most ($n = 287$, 70%) of the sample were born in the UK, and 92% ($n = 380$) lived in the UK.

The Netherlands sample ($n = 278$) had a mean age of 26.6 ($SD = 9.7$), ranging from 18 to 81, and were predominantly educated to pre-bachelor ($n = 137$, 49%) or bachelor ($n = 77$, 28%) levels. The majority ($n = 151$, 54%) were female, with 127 males (46%). Nearly all (266; 96%) of participants were born in the Netherlands, and most ($n = 250$; 90%) were living there.

The Polish sample ($n = 183$) had a mean age of 28.7 ($SD = 11.8$), ranging from 18 to 80, and were most commonly educated to school ($n = 99$, 54%) or Masters level ($n = 47$, 26%). The majority of participants were female ($n = 134$, 74%), with 48 males (26%). Nearly all participants were born ($n = 181$, 99%) and lived ($n = 179$, 98%) in Poland.

Divergence from Preregistration

First, some data was collected before the survey setup satisfied all requirements of the preregistration e.g. timings to allow calculation of RSI. This data has not been included in the current analyses. Second, more data was collected than was targeted to account for the high number of partial completions or because teams had more resources than anticipated. Data analysed in the current manuscript only refers to the stated number of full completions (300/500) per country. These two divergences have led to creation of additional data not analysed in the current manuscript, which may be of benefit to include in any secondary analyses. This supplementary data is available on the OSF page of the project (osf.io/2gsmk).

Pre-Registered Analyses

Data from each world region was analysed separately, and using R. First, Confirmatory Factor Analyses (CFA) were conducted to ensure consistent factor structures. Fit to the data was considered adequate with values of $\leq .08$ for the RMSEA (Browne & Cudeck, 1993), and $\geq .90$ for the CFI and TLI (Bentler & Bonett, 1980) with values above .95 preferred (Hu & Bentler, 1999). These cut-offs were evaluated in conjunction with the standardised factor loadings, as where relevant, slight violations of fit indices cut-offs could be permissible when factor loading is especially high (McNeish, An, & Hancock, 2018). Fit and factor loadings for all models can be found in Table 2.

Table 2: CFA Results

Scale	Country	RMSEA	CFI	TLI	Factor Loadings
4-Factor Humour Model	UK	.059	.852	.840	.492-1.613
	NL	.064	.811	.795	.440-1.339
	PL	.058	.828	.813	.351-1.095
Well-being	UK (-item 2)	.160 (.037)	.935 (.998)	.869 (.993)	.786-1.032 (.844-.988)
	NL	.104	.968	.937	.668-.896
	PL	.092	.973	.946	.687-914
Self-Esteem	UK (items 2 & 3)	.100 (.071)	.939 (.977)	.922 (.967)	.438-.761 (.428-.772)
	NL	.078	.951	.938	.268-.749
	PL (-item 2)	.088 (.064)	.922 (.962)	.899 (.950)	.271-.556 (.270-.560)
Satisfaction with Life	UK	.074	.989	.978	1.190-1.503
	NL	.120	.976	.952	1.138-1.494
	PL	.044	.996	.991	1.007-1.227
Friendship Quality	UK	.083	.962	.952	.604-.844
	NL	.079	.971	.963	.593-.792
	PL (-items 3 & 7)	.120 (.081)	.915 (.970)	.891 (.958)	.510-788 (.508-.791)

Fit indices for the humour data did not meet intended cut-offs, however the removal of items associated with high modification indices led to insufficient fit gains before factor suppression, where all four humour types would not be represented for analysis. As such, no items were removed to maintain the authenticity of the original scale. Minor edits to the well-being, self-esteem and friendship quality scales were required to meet the required standards (see items removed and subsequent results reported in parenthesis in Table 2) and some minor deviations in fit were accepted where removal of further items would have led to over-saturated models.

Second, scale scores for each humour type were generated and standardised into z-scores. As outliers can be problematic for the formulation of clusters (Liu, Li, Wu, & Fu, 2018), individuals with extreme z-scores (>3.29 or <-3.29) were removed to minimise the distortion of profiles. This led to the removal of 4 participants from the UK who had extremely high affiliative humour scores, 2 participants in the Netherlands with extremely high affiliative scores and 2 participants from Poland who had extremely high aggressive humour scores. The mean, standard deviation, Cronbach's alpha and correlations between variables for each country are presented in Tables 1-3 within the Online Supplemental Material.

Third, Cluster Analyses for each world region were conducted to explore both the three and four-profile solutions reported by Galloway (2010), Leist and Müller (2013), Evans and Steptoe-Warren (2018) and Sirigatti et al. (2016). Here, k-means clustering was adopted as the number of expected profiles was known. Table 3 outlines each humour style cluster, noting the number of participants in each, and the mean humour type scores for individuals belonging to each cluster.

Table 3: Mean score of humour types within each humour cluster

UK	Cluster	1	2	3	HHH	1	2	3	4
	Group N	127	126	153		114	60	116	116
	% of total N	31%	31%	38%		28%	15%	29%	29%
	Affiliative	.36	-.91	.56		.04	-1.56	.64	.26
	Aggressive	-.62	-.31	.79		-.06	-.45	.95	-.63
	Self-Enhancing	.35	-.91	.50		-.54	-.92	.86	.20
	Self-Defeating	-.70	-.10	.69		.74	-.60	.48	-.87
<hr/>									
Netherlands	Cluster	1	2	3	HHH	1	2	3	4
	Group N	101	109	66		52	98	70	56
	% of total N	37%	39%	24%		19%	36%	25%	20%
	Affiliative	-.62	.18	.77		-	.57	.21	.22
	Aggressive	-.64	.68	-.11		-	1.46	.25	-1.00
	Self-Enhancing	-.66	-.04	1.11		-.77	.93	-.55	-.20
	Self-Defeating	-.81	.77	-.00		-.55	.32	.80	-1.04
<hr/>									
Poland	Cluster	1	2	3	HHH	1	2	3	4
	Group N	61	58	62		47	41	49	44
	% of total N	34%	32%	34%		26%	23%	27%	24%
	Affiliative	-.50	.62	-.15		.38	.63	.17	-1.25
	Aggressive	.41	.14	-.67		-.74	.12	.72	-.29
	Self-Enhancing	-.72	.97	-.21		.02	1.21	-.44	-.68
	Self-Defeating	.43	.58	-.99		-.81	.63	.65	-.50

Humour style profiles of Galloway (2010), Leist and Müller (2013), Evans and Steptoe-Warren (2018) or Sirigatti et al. (2016) were to be considered replicated in any given country if the profile of each style duplicated those reported. For example, successful replication of Leist and Müller (2013) would require a three-factor solution with profiles representing a) above average use of all humour types, b) below average use of all humour types, and c) greater than average affiliative and self-enhancing humour, and lower than average aggressive and self-defeating humour.

The current analyses only reported one replication: UK data replicated the three-cluster profiles reported by Leist & Müller (2013). All other profiles deviated from those previously reported. Some additional trends are worth noting however. As indicated by positive mean scores on all humour types, the ‘humour endorser’ profile, where individuals use all humour types more than average, was present in all cluster solutions except that of the three-cluster analysis of the Netherlands data. As indicated by negative mean scores on all

humour types, the 'humour denier' profile, where individuals use all humour types below average, was reported in all analyses.

Fourth, to determine the value of such profiles, two sets of regressions to predict psychological health, self-esteem, satisfaction with life and friendship quality were conducted. For the first, the four humour types were entered as continuous predictors (Step 1). For the second set, the humour styles (dummy-coded) were entered (Step 2). Finally, both types and styles were input simultaneously as predictors (Step 3). These analyses, conducted for both three- and four-cluster styles, were adopted to determine whether humour styles provide any additional predictive validity of outcomes over individual types, and vice versa. Combined, these three steps determine the extent to which humour style profiles are valuable for the prediction of outcomes. Given that outcomes are likely related, with all but one representing a measure of psychological health, a conservative Bonferroni correction to the regression models was applied. Thus, the alpha-level was revised to .0125. Regression analyses for three-factor clusters can be found in Table 4, and four-factor clusters can be found in Table 5.

As can be seen from the R^2 values for steps 1 and 2 respectively, the humour types explained much greater proportion of variance in outcomes than humour styles. Typically, humour styles predicted roughly half the variance of that predicted by humour types. Examining the R^2 change between steps 1 and 3, we see that the addition of humour styles adds relatively little to the prediction of outcomes beyond that made by humour types, with the sole exception of the prediction of well-being when looking at the four-clusters found in the UK data. Comparatively, the R^2 change between steps 2 and 3 suggest that humour types provide consistent incremental predictive validity above humour styles across all outcomes, countries and number of clusters. Together, these results suggest humour styles hold relatively little value for the understanding of well-being and friendship, particularly when compared to humour types.

Table 4: Regression Results for 3 Clusters

Country	Outcome	STEP		STEP		STEP		Step 1 -> 3	Step 2 -> 3
		Step 1 Beta	R ² 95% CI	Step 2 Beta	R ² 95% CI	Step 3 Beta	R ² 95% CI	Δ R ² 95% CI	Δ R ² 95% CI
UK	Well-being	AF = .08	.17*	CL1 = .08	.08*	AF = .03	.18*	.01	.09*
		AG = .01	.10, .23	CL2 = -.24*	.04, .14	AG = -.02	.10, .23	-.01, .02	.04, .14
		SE = .33*				SE = .28*			
		SD = -.24*				SD = -.27*			
	Self-esteem	CL1 = -.09		CL1 = .27*	.11*	AF = .14	.36*	.01	.25*
		CL2 = -.15		CL2 = -.09	.05, .16	AG = .09	.28, .41	-.01, .02	.18, .32
				SE = .40*		SE = .35*			
				SD = -.48*		SD = -.50*			
	Satisfaction with life	CL1 = -.06		CL1 = .22*	.06*	AF = .10	.13*	.00	.06*
		CL2 = .14		CL2 = -.06	.02, .11	AG = -.01	.06, .17	-.01, .01	.02, .11
				SE = .17*		SE = .17*			
				SD = -.27*		SD = -.29			
	Friendship quality	CL1 = .06		CL1 = .12	.04*	AF = .31*	.14*	.01	.10*
		CL2 = .02		CL2 = -.11	.01, .08	AG = -.13	.07, .19	-.01, .03	.05, .15
				SE = .21*		SE = .14*			
				SD = -.24*		SD = -.20*			
NL	Well-being	CL1 = -.07		CL1 = -.21*	.05*	AF = .15	.13*	.01	.08*
		CL2 = .11		CL2 = -.28*	.01, .11	AG = .13	.05, .19	-.01, .03	.02, .14
				SE = .21*		SE = .22*			
				SD = -.17		SD = -.24*			
		CL1 = .05				CL1 = .05			
		CL2 = -.11				CL2 = -.11			

CONSISTENCY AND VALUE OF HUMOUR STYLES

15

	Self-esteem	AF = .17* AG = .11 SE = .29* SD = -.42*	.26* .17, .34	CL1 = -.23* CL2 = -.35*	.08* .02, .14	AF = .18* AG = .13 SE = .28* SD = -.40* CL1 = .02 CL2 = -.04	.26* .16, .33	.00 -.01, .01	.19* .11, .27
	Satisfaction with life	AF = .14 AG = .09 SE = .23* SD = -.22*	.13* .05, .19	C1 = -.28* CL2 = -.31*	.07* .02, .13	AF = .14 AG = .14 SE = .18 SD = -.16 CL1 = -.06 CL2 = -.17	.14* .06, .20	.01 -.01, .03	.07* .01, .13
	Friendship quality	AF = .25* AG = -.04 SE = .15 SD = -.25*	.15* .07, .21	CL1 = -.22* CL2 = -.29*	.06* .01, .11	AF = .24* AG = -.05 SE = .13 SD = -.26* CL1 = -.06 CL2 = -.03	.15* .06, .21	.00 -.01, .01	.09* .03, .16
Poland	Well-being	AF = .13 AG = -.09 SE = .37* SD = -.14	.20* .09, .28	CL1 = -.27* CL2 = .07	.10* .02, .18	AF = .15 AG = -.07 SE = .42* SD = -.09 CL1 = -.05 CL2 = -.13	.20* .08, .28	.01 -.01, .02	.11* .03, .19
	Self-esteem	AF = .08 AG = -.04 SE = .44* SD = -.32*	.28* .16, .37	CL1 = -.33* CL2 = .96	.13* .05, .22	AF = .08 AG = -.06 SE = .42* SD = -.36* CL1 = .06 CL2 = .08	.28* .15, .36	.00 -.01, .01	.15* .06, .24
	Satisfaction with life	AF = .00 AG = -.06 SE = .43* SD = -.25*	.22* .11, .31	CL1 = -.35* CL2 = -.02	.11* .04, .20	AF = .00 AG = -.03 SE = .45* SD = -.18 CL1 = -.10 CL2 = -.12	.22* .10, .30	.01 -.01, .02	.11* .03, .19

CONSISTENCY AND VALUE OF HUMOUR STYLES

17

CL3 = -.04

	Friendship quality	AF = .23* AG = -.10 SE = .14* SD = -.20*	.13 .06, .18	CL1 = -.23* CL2 = -.20* CL3 = -.02	.06* .02, .11	AF = .24* AG = -.17* SE = .10 SD = -.23* CL1 = .05 CL2 = .04 CL3 = .17	.14* .07, .19	.01 -.01, .03	.07* .03, .12
NL	Well-being	AF = .13 AG = .07 SE = .22* SD = -.24*	.12* .05, .19	CL1 = -.17 CL2 = .05 CL3 = -.24*	.08* .03, .14	AF = .14 AG = .07 SE = .17 SD = -.20 CL1 = -.02 CL2 = .02 CL3 = -.09	.13* .04, .18	.01 -.01, .02	.04* .00, .09
	Self-esteem	AF = .17* AG = .11 SE = .29* SD = -.42*	.26* .17, .34	CL1 = -.20* CL2 = .03 CL3 = -.28*	.10* .04, .16	AF = .18 AG = .08 SE = .28* SD = -.46* CL1 = .05 CL2 = .08 CL3 = .09	.27* .16, .33	.00 -.01, .01	.16* .09, .24
	Satisfaction with life	AF = .14 AG = .09 SE = .23* SD = -.22*	.13* .05, .19	CL1 = -.15 CL2 = .08 CL3 = -.19	.07* .02, .13	AF = .17 AG = .10 SE = .21 SD = -.19 CL1 = .03 CL2 = -.01 CL3 = -.06	.13* .05, .19	.00 -.01, .01	.06* .01, .11
	Friendship quality	AF = .25* AG = -.04 SE = .15 SD = -.25*	.15* .07, .21	CL1 = -.26* CL2 = -.08 CL3 = -.24*	.06* .01, .12	AF = .28* AG = -.03 SE = .19 SD = -.25* CL1 = .03 CL2 = -.05	.15* .06, .21	.00 -.01, .01	.09* .02, .15

CONSISTENCY AND VALUE OF HUMOUR STYLES

18

						CL3 = .01			
PL	Well-being	AF = .13	.20*	CL1 = .41*	.15*	AF = -.01	.23*	.03	.08*
		AG = -.09	.09, .28	CL2 = .35*	.05, .23	AG = -.07	.10, .31	-.01, .08	.01, .15
		SE = .37*		CL3 = .14		SE = .39*			
		SD = -.14				SD = -.13			
						CL1 = .26			
						CL2 = .12			
						CL3 = .20			
	Self-esteem	AF = .08	.28*	CL1 = .26*	.08*	AF = .11	.30*	.02	.22*
		AG = -.04	.16, .37	CL2 = .19	.01, .15	AG = -.08	.17, .38	-.01, .05	.12, .32
		SE = .44*		CL3 = -.02		SE = .56*			
		SD = -.32*				SD = -.32*			
						CL1 = -.06			
						CL2 = -.17			
						CL3 = .05			
	Satisfaction with life	AF = .00	.22*	CL1 = .18	.05*	AF = .00	.24*	.02	.19*
		AG = -.06	.11, .31	CL2 = .17	.00, .11	AG = -.13	.11, .31	-.02, .06	.09, .29
		SE = .43*		CL3 = -.02		SE = .53*			
		SD = -.25*				SD = -.29*			
						CL1 = -.05			
						CL2 = -.09			
						CL3 = .13			
	Friendship quality	AF = .18	.15*	CL1 = .30*	.08*	AF = .26*	.16*	.01	.08*
		AG = -.13	.06, .24	CL2 = .26*	.01, .15	AG = -.11	.05, .23	-.02, .03	.01, .15
		SE = .22*		CL3 = .10		SE = .23			
		SD = -.20*				SD = -.16			
						CL1 = .04			
						CL2 = .02			
						CL3 = .15			

Note: CL1 = Cluster 1; CL2 = Cluster 2; CL3 = Cluster 3; AFF = Affiliative Humour; AGG = Aggressive Humour; SE = Self-Enhancing Humour; SD = Self-Defeating Humour; * = p<.01

Discussion

Cluster Replication

Using data collected from the UK, Netherlands, and Poland, the current study first explored humour style profiles. Findings from several sets of cluster analyses suggest there is moderate consistency in humour style profiles across countries. Evident within both the three- and four-cluster analyses from all countries was the 'humour denier' profile (below-average use of all humour types), which has been inconsistently identified throughout the extant literature (see Table 1). Evident within all humour style analyses, except that from the three-cluster Netherlands data, was also the consistently reported 'humour endorser' profile (above-average use of all humour types). Finally, the more positive humour user profile (above average affiliative and self-enhancing with below-average aggressive and self-defeating humour use) was only identified in some analyses despite being consistently reported across the extant literature. The remaining humour styles identified had diverse profiles and provided an account of humour use inconsistent across countries and with those reported within the extant literature.

Current analyses reported only one full replication of previous humour style profiles as identified by the existing literature. The UK data replicated the three-cluster profiles reported by Leist & Müller (2013). Here profiles representing the humour endorser, humour denier, and positive humour user (as defined above) were replicated. All other analyses evidenced at least one profile which deviated and thus failed to replicate the styles of humour use previously reported. That is, the current study failed to replicate the humour style profiles previously reported by Galloway (2010), Evans and Steptoe-Warren (2018) and Sirigatti et al. (2016).

In sum, the current study reported limited consistency in the identification of humour style profiles, despite focusing upon adult populations only, adopting identical analytical processes, and capturing reasonable sample sizes. As we have not compared two like-for-like countries, the current research therefore provides initial evidence for one likely interpretation: cultural differences in humour (e.g. Chen & Martin, 2007; Martin & Sullivan, 2013; Yue, Jiang, Lu, & Hiranandani, 2016) which may lead to substantively different patterns of humour use. The extent of cultural differences is yet unclear however it may make comparisons of humour styles across cultures problematic. The current findings conclude that the consistency of humour style profiles is limited, possibly due to cultural differences.

Prediction of Outcomes

The second aim of the current study was to examine the value of humour styles for the prediction of friendship quality and three indicators of well-being. Here, the results suggest that humour types consistently predict much greater proportions of variance in relevant outcomes than predicted by humour styles. Humour styles typically predicted about half of the variance in outcomes predicted by humour types, and provided negligible incremental predictive validity over them. These results were consistent across countries and across the two sets of analyses considering three and four humour styles. Such results directly contradict those of Leist & Müller (2013) who used dichotomised humour type scores to evidence the value of humour clusters, and provides evidence to suggest the initial promise of grouping individuals based upon their styles of humour use may have been exaggerated.

In sum, the current study reported consistent support for the claim that humour styles hold limited value for the prediction of important humour outcomes, particularly when compared to humour types. This set of findings further questions the purpose of humour styles, particularly as they suggest that even a successful

replication of style profiles do not equate to greater predictive value than that presented by considering the four humour types which informed them. As such, the current study provides no support for encouraging application of such cluster analysis findings to practice. Instead, the current study argues for greater emphasis to be placed upon more closely considering the individual humour types and any possible interactions between them.

Limitations and Future Directions

Findings of the current study should be interpreted within context of its limitations. First, the current study did not provide an opportunity for any direct replication of previous findings reported (Table 1) as no country represented in the extant literature was captured through the current data collection. In the closest similarity, Evans and Steptoe-Warren (2018) also collected UK data but their use of an other-report measurement strategy may have introduced differences in ratings of humour which make comparisons inappropriate (Fine, 1975). Indeed, the only successful replication was that of UK data equivalent to styles first reported from a German sample (Leist & Müller, 2013). Caution is therefore required before drawing conclusive statements on the consistency of humour styles, and researchers are encouraged to explore within-culture consistency using a similar pre-registered design, either testing the countries already explored or by collecting two independent samples and comparing profiles.

Second, the current study had modest sample sizes following the exclusion criteria and was limited to considering three/four styles of humour. Clustering is highly sensitive to sample size, so development of more nuanced methods for planning sample size analogous to power analysis represents a valuable investment for the field to support more robust clustering. Larger sample sizes are of particular benefit as, in the endeavour to identify a small number of consistent humour profiles across the extant literature, there has been insufficient sample size and researcher inclination to consider or explore larger numbers of humour styles. Given the complexity and multi-dimensionality of humour, it would not be unreasonable to consider that there are a much greater number of humour styles which may be of value to study. The ongoing dominance of three/four cluster analyses may have consequences for interpretations of the consistency and value of humour styles because cluster analysis will always create groups regardless of sample size or meaningfulness of grouping. As such, it is possible that the styles identified can represent statistical artefacts more than meaningful groups. Research with much larger samples therefore seem to be of benefit for facilitating examination of a greater range of styles and more robust evaluations of their consistency and value. Should cluster analysis remain problematic as an analytical approach to establishing consistent and meaningful groups, alternative strategies to summarise these complex phenomena should be explored.

Whilst not the primary focus of the current study, encouraging caution in grouping-based analyses across different fields, and encouraging more robust evaluation of such practices, is to be encouraged. For example, there are a number of very popular occupational recruitment and training measures which aim to group individuals into different categories based upon their traits, but that these are often supported with limited evidence (e.g. Furnham, 2017). Similar concerns surrounding grouping can be seen with clinical diagnoses (Allsopp et al., 2019) and educational practices such as learning styles (Kirschner, 2017). These types of analyses seem particularly susceptible to questionable research practices, from determining how many groups to report, criteria of group membership, and method used to evaluate their value. Thus, greater adoption of more robust open science behaviours in these practices would be of significant value. In particular, greater use of open data to facilitate secondary analyses of alternative interpretations, and pre-

registration and/or registered reports to minimise questionable opportunistic practices, seems of particular benefit to limit researcher degrees of freedom (Munafò et al., 2017).

Finally, humour is just one of many behavioural strategies considered within emotion regulation models (e.g. Samson & Gross, 2012). Within this literature there has been growing differentiation between types and styles (e.g. Hampton et al., 2015) and between the identification, selection and implementation of regulation strategies (Gross, 2015). Such differentiations may be complementary and of benefit in driving more nuanced models of humour use. The current study provides modest additional evidence to encourage scrutiny of the development of emotion regulation styles, particularly where numerous emotion regulation strategies are retrospectively reported and then statistically combined into groups with little theoretical justification. In-line with recent recommendations within the personality literature to focus upon the specific sub-facets of interest (de Vries, de Vries & Born, 2011), the current study provides stronger grounds for recognising the nuances behind each specific humour type, and engaging more with the cultural and contextual factors which influence their selection, implementation and interpretation.

Conclusions

The current study explored the consistency in humour styles profiles across the UK, Netherlands and Poland, and their value for the prediction of friendship quality and well-being. Humour styles identified were somewhat inconsistent across countries and when compared to the existing literature. Furthermore, humour styles held little value for the prediction of relevant outcomes. The importance of humour style clusters is therefore undermined by a) the inconsistency with which they can be consistently formed, possibly due to cultural differences, and b) their limited value in predicting outcomes when compared with humour types. Future research in this field must continue to minimise researcher degrees of freedom through pre-registration, look to achieve sample sizes capable of exploring a much larger range of humour styles, and/or consider alternative theoretical and analytical strategies to inform understanding of the experiences and consequences of humour.

Open Science Statement

The current study was pre-registered (<https://osf.io/9kjp8>) and has been published as a Registered Report. Ethical approval, study materials, raw (anonymised), summary and meta-data, and analysis code, are all publicly accessible on the Open Science Framework (<https://osf.io/2gsmk>).

Contributions (CreDiT)

Responsibility	Initials
Conceptualization	TRE, GSW, SEJ, GG
Data curation	TRE, NJ, GN
Formal analysis	TRE, NJ
Funding acquisition	-
Investigation	TRE, SEJ, NJT, DF, AvS, IMM, LD, AGN, NJ, AS, JW
Methodology	TRE
Project administration	TRE
Resources	TRE, AvS, AGN, JW, NJ
Software	TRE, NJ
Supervision	TRE
Validation	TRE, NJ, GN
Visualization	-
Writing - original draft	TRE
Writing – review & editing	TRE, LD, DF, GG, AGN, SEJ, GN, AS, AvS, GSW, NJT, IMM, NDU, NJ, JW

References

- Allsopp, K., Read, J., Corcoran, R., & Kinderman, P. (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry Research*, *279*, 15-22.
- Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence*, *16*(5), 427-454.
- Atienza, F. L., Balaguer, I., & García-Merita, M. L. (2003). Satisfaction with life scale: Analysis of factorial invariance across sexes. *Personality and Individual Differences*, *35*(6), 1255-1260.
- Bariaud, F. (1989). Age differences in children's humor. *Journal of Children in Contemporary Society*, *20*(1-2), 15-45.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606.
- Billig, M. (2005). *Laughter and ridicule: Towards a social critique of humour*. Great Britain: Sage.
- Blom, E. H., Bech, P., Högberg, G., Larsson, J. O., & Serlachius, E. (2012). Screening for depressed mood in an adolescent psychiatric context by brief self-assessment scales—testing psychometric validity of WHO-5 and BDI-6 indices by latent trait analyses. *Health and Quality of Life Outcomes*, *10*(1), 149.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-cultural Psychology*, *1*(3), 185-216.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230-258.
- Chang, J. H., Chen, H. C., Hsu, C. C., Chan, Y. C., & Chang, Y. L. (2015). Flexible humor styles and the creative mind: Using a typological approach to investigate the relationship between humor styles and creativity. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 306-312.
- Chen, G. H., & Martin, R. A. (2007). A comparison of humor styles, coping humor, and mental health between Chinese and Canadian university students. *Humor*, *20*(3), 215-234.
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71-75.
- Dolnicar, S., Grün, B., Leisch, F., & Schmidt, K. (2014). Required sample sizes for data-driven market segmentation analyses in tourism. *Journal of Travel Research*, *53*(3), 296-306.
- Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science*, *37*(2), 191-203.
- Evans, T. R., & Steptoe-Warren, G. (2018). Humor style clusters: Exploring managerial humor. *International Journal of Business Communication*, *55*(4), 443-454.
- Fine, G. A. (1975). Components of perceived sense of humor ratings of self and other. *Psychological Reports*, *36*(3), 793-794.
- Ford, T. E., McCreight, K. A., & Richardson, K. (2014). Affective style, humor styles and happiness. *Europe's Journal of Psychology*, *10*(3), 451-463.
- Ford, T. E., Lappi, S. K., O'Connor, E. C., & Banos, N. C. (2017). Manipulating humor styles: Engaging in self-enhancing humor reduces state anxiety. *Humor*, *30*(2), 169-191.
- Formann, A. K. (1984). *Die latent-class-analyse: Einführung in Theorie und Anwendung*. Beltz.
- Fox, C. L., Hunter, S. C., & Jones, S. E. (2016). Children's humor types and psychosocial adjustment. *Personality and Individual Differences*, *89*, 86-91.
- Furnham, A. (2017). *Myers-Briggs Indicator (MBTI)*. In Virgil Zeigler-Hill & Todd K. Shackelford (Eds). *The SAGE Handbook of personality and individual differences*. New York: Sage.
- Galloway, G. (2010). Individual differences in personal humor styles: Identification of prominent patterns and their associates. *Personality and Individual Differences*, *48*(5), 563-567.
- Gorrese, A., and Ruggieri, R. (2012). Peer attachment: A meta-analytic review of gender and age differences and associations with parent attachment. *Journal of Youth and Adolescence*, *41*(5), 650-672.
- Greengross, G. (2013). Humor and aging—a mini-review. *Gerontology*, *59*(5), 448-453.

- Gross, J. J. (2015). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry*, 26(1), 130-137.
- Hajos, T. R., Pouwer, F., Skovlund, S. E., Den Oudsten, B. L., Geelhoed-Duijvestijn, P. H. L. M., Tack, C. J., & Snoek, F. J. (2013). Psychometric and screening properties of the WHO-5 well-being index in adult outpatients with Type 1 or Type 2 diabetes mellitus. *Diabetic Medicine*, 30(2), e63-e69.
- Hall, J. A. (2017). Humor in romantic relationships: A meta-analysis. *Personal Relationships*, 24(2), 306-322.
- Hampton, A. J., Hadjistavropoulos, T., Gagnon, M. M., Williams, J., and Clark, D. (2015) 'The Effects of Emotion Regulation Strategies on the Pain Experience: A Structured Laboratory Investigation'. *Pain*, 156 (5), 868-879
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Huang, K. P., & Kuo, W. C. (2011). Does humor matter? From organization management perspective. *Journal of Social Sciences*, 7(2), 141-145.
- Hughes, D. J., & Evans, T. R. (2016). Comment: Trait EI moderates the relationship between ability EI and emotion regulation. *Emotion Review*, 8(4), 331-332.
- Kaczmarek, L. D., Bujacz, A., & Eid, M. (2015). Comparative latent state-trait analysis of satisfaction with life measures: The steen happiness index and the satisfaction with life scale. *Journal of Happiness Studies*, 16(2), 443-453.
- Kirschner, P. A. (2017). Stop propagating the learning styles myth. *Computers & Education*, 106, 166-171.
- Kuiper, N. A., Martin, R. A., & Olinger, L. J. (1993). Coping humour, stress, and cognitive appraisals. *Canadian Journal of Behavioural Science*, 25(1), 81-96.
- Kuiper, N., Kirsh, G., & Maiolino, N. (2016). Identity and intimacy development, humor styles, and psychological well-being. *Identity*, 16(2), 115-125.
- Leiner, D. J. (2013). Too fast, too straight, too weird: Post hoc identification of meaningless data in internet surveys. Working Paper 11/30/2013, 1-34. Retrieved from https://www.researchgate.net/publication/258997762_Too_Fast_too_Straight_too_Weird_Post_Hoc_Identification_of_Meaningless_Data_in_Internet_Surveys
- Leist, A. K., & Müller, D. (2013). Humor types show different patterns of self-regulation, self-esteem, and well-being. *Journal of Happiness Studies*, 14(2), 551-569.
- Li, S., & Seale, C. (2007). Managing criticism in Ph. D. supervision: A qualitative case study. *Studies in Higher Education*, 32(4), 511-526.
- Liu, H., Li, J., Wu, Y., & Fu, Y. (2018). Clustering with outlier removal. arXiv preprint arXiv:1801.01899.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19-40.
- Martin, G. N., & Sullivan, E. (2013). Sense of humor across cultures: A comparison of British, Australian and American respondents. *North American Journal of Psychology*, 15(2), 375-384.
- Martin, R. A. (1996). The situational humor response questionnaire (SHRQ) and coping humor scale (CHS): A decade of research findings. *Humor*, 9(3-4), 251-272.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the humor styles questionnaire. *Journal of Research in Personality*, 37(1), 48-75.
- Martin, R. A. & Ford, T. (2018). *The psychology of humor: An integrative approach*. USA: Elsevier.
- McCosker, B., & Moran, C. C. (2012). Differential effects of self-esteem and interpersonal competence on humor styles. *Psychology Research and Behavior Management*, 5, 143-150.
- McNeish, D., An, J., & Hancock, G. R. (2018). The thorny relation between measurement quality and fit index cutoffs in latent variable models. *Journal of Personality Assessment*, 100(1), 43-52.
- Mesmer-Magnus, J., Glew, D. J., & Viswesvaran, C. (2012). A meta-analysis of positive humor in the workplace. *Journal of Managerial Psychology*, 27(2), 155-190.
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie, N., ... Wagenmakers, E.-J. (2017). A manifesto for reproducible science. *Human Nature Behavior*, 1, 1-9.

- Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology, 3*(2), 137-152.
- Robert, C., & Yan, W. (2007). *The case for developing new research on humor and culture in organizations: Toward a higher grade of manure*. In J. J. Martocchio, J. J. Martocchio (Eds.), *Research in personnel and human resources management*, Vol 26 (pp. 205-267). US: Elsevier Science/JAI Press.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1-25.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Samson, A. C., & Gross, J. J. (2012). Humour as emotion regulation: The differential consequences of negative versus positive humour. *Cognition & Emotion, 26*(2), 375-384.
- Sarstedt, M., & Mooi, E. (2014). *Cluster analysis*. In M. Sarstedt and E. Mooi (Eds.), *A concise guide to market research* (pp. 273-324). Berlin: Springer.
- Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology, 89*(4), 623-642.
- Schneider, M., Voracek, M., & Tran, U. S. (2018). "A joke a day keeps the doctor away?" Meta-analytical evidence of differential associations of habitual humor styles with mental health. *Scandinavian Journal of Psychology, 59*(3), 289-300.
- Schnurr, S., & Chan, A. (2009). Politeness and leadership discourse in New Zealand and Hong Kong: A cross-cultural case study of workplace talk. *Journal of Politeness Research, 5*(2), 131-157.
- Sirigatti, S., Penzo, I., Giannetti, E., Casale, S., & Stefanile, C. (2016). Relationships between humorism profiles and psychological well-being. *Personality and Individual Differences, 90*, 219-224.
- Stieger, S., Formann, A. K., & Burger, C. (2011). Humor styles and their relationship to explicit and implicit self-esteem. *Personality and Individual Differences, 50*(5), 747-750.
- Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 well-being index: a systematic review of the literature. *Psychotherapy and psychosomatics, 84*(3), 167-176.
- Tucker, R. P., Wingate, L. R., O'Keefe, V. M., Sligh, M. L., Judah, M. R., & Rhoades-Kerswill, S. (2013). The moderating effect of humor style on the relationship between interpersonal predictors of suicide and suicidal ideation. *Personality and Individual Differences, 54*(5), 610-615.
- de Vries, A., de Vries, R. E., & Born, M. P. (2011). Broad versus narrow traits: Conscientiousness and Honesty-Humility as predictors of academic criteria. *European Journal of Personality, 25*(5), 336-348.
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. a. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science, 7*, 627-633.
- Wanzer, M. B., Wojtaszczyk, A. M., & Kelly, J. (2009). Nurses' perceptions of physicians' communication: the relationship among communication practices, satisfaction, and collaboration. *Health Communication, 24*(8), 683-691.
- Whisman, M. A., & Judd, C. M. (2016). A cross-national analysis of measurement invariance of the satisfaction with life scale. *Psychological Assessment, 28*(2), 239.
- Wilkinson, R. B., & Goh, D. Y. L. (2014). Structural, age, and sex differences for a short form of the inventory of parent and peer attachment: The IPPA-45. *Journal of Relationships Research, 5*, e5.
- Wilson, J. M., & Wilkinson, R. B. (2012). The self-report assessment of adolescent attachment: A systematic review and critique. *Journal of Relationships Research, 3*, 81-94.
- World Health Organization: Regional Office for Europe (1998) *Well-Being measures in primary health care: The DepCare Project*. Consensus meeting, Stockholm.
- Yip, J. A., & Martin, R. A. (2006). Sense of humor, emotional intelligence, and social competence. *Journal of Research in Personality, 40*(6), 1202-1208.
- Yue, X., Jiang, F., Lu, S., & Hiranandani, N. (2016). To be or not to be humorous? Cross cultural perspectives on humor. *Frontiers in Psychology, 7*, 1495.
- Yue, X. D., Liu, K. W. Y., Jiang, F., & Hiranandani, N. A. (2014). Humor styles, self-esteem, and subjective happiness. *Psychological Reports, 115*(2), 517-525.

Zeigler-Hill, V., McCabe, G. A., & Vrabel, J. K. (2016). The dark side of humor: DSM-5 pathological personality traits and humor styles. *Europe's Journal of Psychology, 12*(3), 363-376.