COUNTING
WHAT
COUNTS.

IMPACT & INSIGHT TOOLKIT: SENTIMENT FRAMEWORK RESEARCH REPORT

RECOMMENDATION REPORT IDENTIFYING EXISTING SENTIMENT FRAMEWORKS SUITABLE FOR THE IMPACT & INSIGHT TOOLKIT

Prepared by Counting What Counts

25th August 2021

https://impactandinsight.co.uk/

Introduction

Background

This report aims to identify and recommend existing sentiment frameworks suitable for the Impact & Insight Toolkit (Toolkit). The report has been conducted in response to National Portfolio Organisation (NPO) feedback from Counting What Counts' (CWC) Artform & Museum Metrics Strand (AMMS). NPOs that took part in AMMS workshops requested new sentiment dimensions to capture overall feelings among audiences and peers. Furthermore, in individual conversations with NPOs, as part of AMMS, it was suggested that the Toolkit needs to capture a breadth of emotions, including negative feelings.

Identifying an appropriate sentiment framework will allow CWC to implement a new set of sentiment dimensions within the Toolkit. The purpose of implementing sentiment dimensions is to: complement and enrich existing dimensions, allowing arts organisations to gain more insight from evaluations; benefit the art sector as a whole by contributing to an anonymous aggregate dataset.

As such the report will evaluate sentiment framework options against two criteria:

- How complete the framework is (i.e. does it encompass positive and negative emotions).
- Suitability to be adapted to Toolkit dimensions.

Approach

Sentiment frameworks are intrinsically linked to emotion classification. Emotion classification is a means to distinguish or categorise emotions. There are two dominant viewpoints within the field of emotion classification: that emotions are discrete and fundamentally different constructs (discrete models); that emotions can be characterised on a dimensional basis in groupings (dimensional models).

This report will identify four emotion classifications that could form the basis for a suitable sentiment framework. It will focus on emotion classifications which utilise dimensional and hybrid models; these models are widely employed within psychology and market research as they offer more complete sentiment frameworks.

The report will start with Plutchik's Model, it will then go on to detail three emotion classifications, the Circumplex Model, the PAD Model and the PANA Model, that are closely linked to methods of self-reporting sentiment, also known as affect measures.

Affect measures are used to study human affect (including emotions and mood) and refer to measures obtained from self-reporting studies asking participants to quantify their current feelings. Affect measures tend to be based on a specific approach to emotion classification. This report will look at the following affect measures that use self-reporting surveys: the Affective Slider (AS), the Self-assessment Manikin (SAM), the PANAS Scale, the PANAVA-KS Scale and the LE-PANAVA Scale.

Research Methods

The report has been compiled using desk-based online research, working primarily with the browser Chrome and Google search engine. All information and examples contained within the report are publicly available. Where academic papers have been accessed online, previews and extracts have been used where the full text is unavailable.

Description of Options

1. Plutchik's Model

The Plutchik Model was initially introduced by Robert Plutchik in 1958, and subsequently developed into Plutchik's Wheel of Emotions in 1980. The model, a hybrid of discrete and dimensional theories, considered there to be eight primary emotions: anger, fear, sadness, disgust, surprise, anticipation, trust and joy. His wheel of emotions was developed to visually describe how these emotions relate. The wheel presents an additional twenty-four 'Primary, 'Secondary' and 'Tertiary' dyads (feelings composed of two emotions).
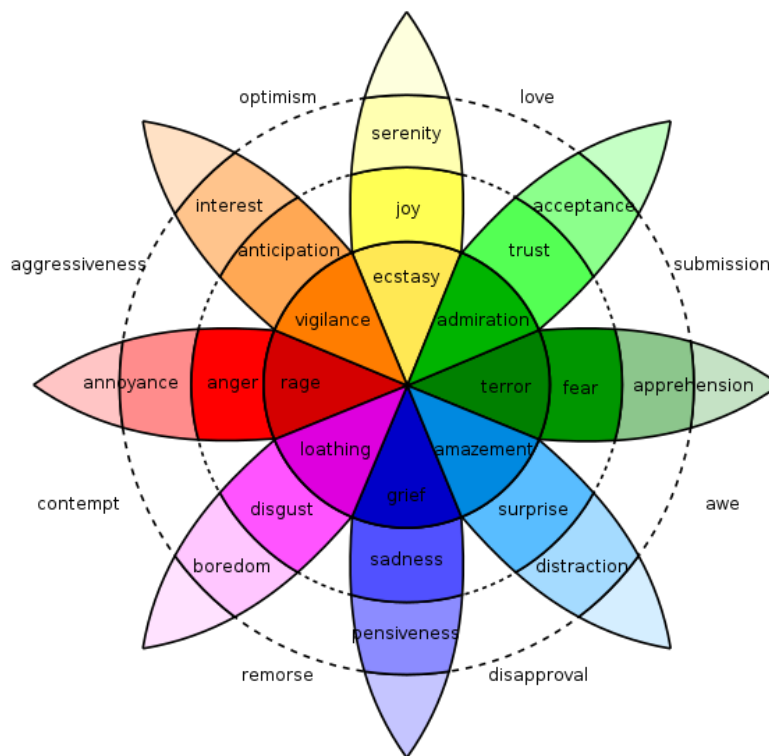

Figure 1: Robert Plutchik's Wheel of Emotions

The benefit of this option is that it presents a complete sentiment framework that encompasses nuanced positive and negative emotions. However, as this is an emotion classification, rather than an affect measure, it would need to be adapted to the self-reporting survey format. This could be achieved by adapting the primary emotions or additional dyads into sentiment dimensions. A limitation of this option is that not all language used in the model is appropriate for Toolkit users.

2. Circumplex Model of Emotion

In contrast to Plutchik's Model, the Circumplex Model of Emotion, developed by James Russell in 1980, maps emotions along x and y axis in a two-dimensional circular space. Based on the theory of multi-dimensional emotions, arousal (activation vs deactivation) represents the vertical axis and valence (unpleasant vs pleasant) represents the horizontal axis, while the centre represents neutral. Emotional responses are plotted on the circumplex according to their levels of arousal and valence.
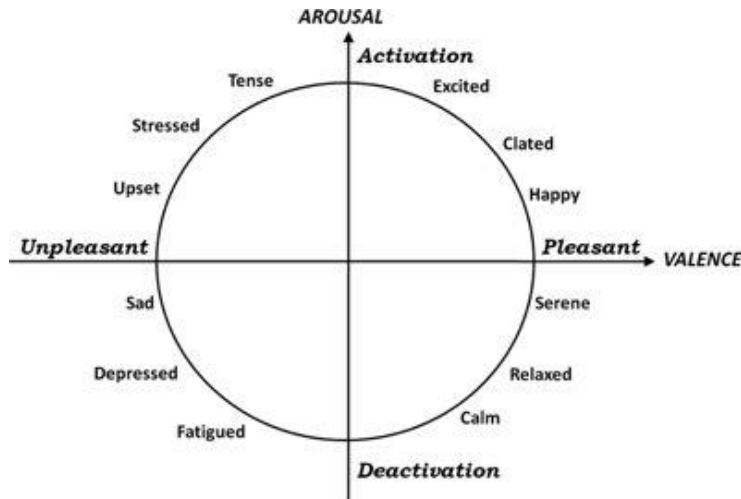

Figure 2: Circumplex Model of Emotion

The Affective Slider (AS)

The Affective Slider (AS), developed by Aberto Betella and Paul Verschure in 2016, is a language free affect measure based on the Circumplex Model. Composed of two slider controls for the quick assessment of arousal and valence, AS was created as a contemporary alternative to the more widely used Self-assessment Manikin (SAM).
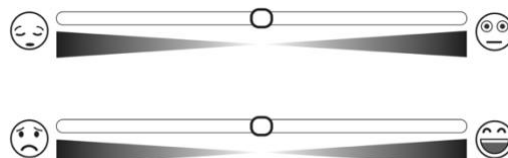

Figure 4: The Affective Slider (AS)

The benefit of the Circumplex Model is that it presents a way to measure a range of positive and negative emotions using two dimensions: arousal and valence. A limitation is that dimensions scores need to be plotted onto the Circumplex Model to be 'read' by Toolkit users. In addition, the language is not appropriate for Toolkit users. However, AS presents a non-verbal, digital means to measure these dimensions.

3. PAD Emotional State Model

The PAD Emotional State Model, developed by Albert Mehrabian and James Russell in 1974, uses three dimensions to represent all emotions. The PAD dimensions are: pleasure, arousal and dominance. PAD is a development of the Circumplex Model of Emotion. The PAD Model is commonly used in market research to study consumer behaviour.
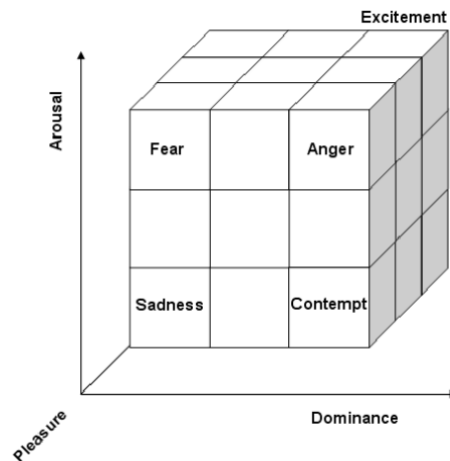

Figure 3: PAD Emotional State Model

Self-assessment Manikin (SAM)

The Self-assessment Manikin (SAM), developed by Peter Lang in 1980, is a language-free affect measure based on the PAD Model. It consists of three rows of pictograms, each of which use stylised figures to capture one dimension on a five-point scale. The use of SAM is widespread within phycology and market research.
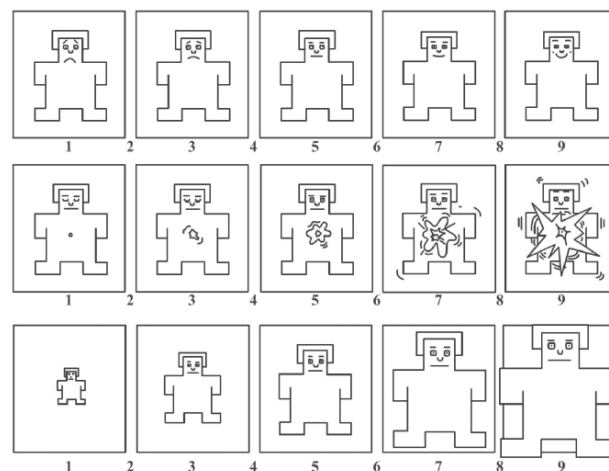

Figure 4: Self-Assessment Manikin

The benefit of the PAD Model is that it presents a way to measure a range of positive and negative emotions using three dimensions: pleasure, arousal, dominance. The addition of a third dimension, dominance, enables the model to capture a more nuanced range of emotions than the Circumplex Model. A limitation is that the language is not appropriate for Toolkit users. However, SAM offers a widely used, quick, accessible, engaging, less intrusive approach that minimises survey fatigue. A limitation of SAM is that it's visual design may seem dated to users.

Much like the Circumplex Model, the PAD Model does not measure specific emotions. However, scores can be used to plot emotions by using, for example, the Semantic Differential Scale devised by Albert Mehrabian and James Russell in 1974.

Table 1

Factor Loadings of Each of the 18 Bipolar Adjective Pairs in the Semantic Differential for Picture Ratings

	Factor 1 "Pleasure"	Factor 2 "Arousal"	Factor 3 "Dominance"
Unhappy–Happy	0.914	0.063	0.148
Annoyed–Pleased	0.883	0.068	0.158
Unsatisfied–Satisfied	0.868	0.144	0.114
Melancholic–Contented	0.725	0.095	0.056
Despairing–Hopeful	0.858	0.063	0.078
Bored–Relaxed	0.580	0.372	0.234
Relaxed–Stimulated	–0.211	0.774	0.052
Calm–Excited	–0.181	0.793	0.056
Sluggish–Frenzied	0.268	0.771	0.005
Dull–Jittery	–0.211	0.793	0.121
Sleepy–Wide awake	–0.046	0.810	0.047
Unaroused–Aroused	0.051	0.827	0.127
Controlled–Controlling	0.262	0.192	–0.673
Influenced–Influential	0.292	0.089	–0.618
Cared for–In control	–0.090	0.198	–0.626
Awed–Important	0.199	–0.040	–0.301
Submissive–Dominant	0.195	0.306	–0.695
Guided–Autonomous	0.161	–0.100	–0.479
Amount of variance accounted for:	24.6	23.12	12.18

Figure 5: Semantic Differential Scale

4. PANA Model

The Positive Activation – Negative Activation (PANA) Model, developed by Auke Tellegen and David Watson in 1985, is based on the idea that positive and negative affect are two separate systems.

Positive affect	Negative affect
Attentive	Hostile
Active	Irritable
Alert	Ashamed
Excited	Guilty
Enthusiastic	Distressed
Determined	Upset
Inspired	Scared
Proud	Afraid
Interested	Jittery
Strong	Nervous

Figure 6: PANA Model

PANAS Scale

The Positive and Negative Affect Schedule (PANAS) Scale, developed from the PANA Model by Lee Clark, Auke Tellegen and David Watson in 1988, is a widely used self-report measure for capturing emotions. The PANAS Scale consists of twenty words, ten negative and ten positive, which users self-report on a five-point scale. The gathered data is used to create a final score indicating high positive affectivity, low positive affectivity, low negative affectivity or high negative affectivity.

A benefit of the PANA model and PANAS Scale is that they recognise that positive and negative emotions can be experienced at the same time. The framework is widely used to capture emotion, however, it doesn't initially present itself as easily adaptable to Toolkit use due to it consisting of twenty individual items. A subset of these items would need to be selected for the Toolkit to minimise survey fatigue.

A limitation of the PANA Model and PANAS Scale is that they do not include many common positive emotion words such as 'happy' or 'joyful', but instead include items such as 'alert', 'active', and 'attentive' that may not always be positive in meaning. Similarly, the measure of negative affect does not include many common negative emotion words such as 'sad' and 'angry'.

	Very slightly/ Not at all	A little	Moderately	Quite a bit	Extremely
Interested	1	2	3	4	5
Alert	1	2	3	4	5
Attentive	1	2	3	4	5
Excited	1	2	3	4	5
Enthusiastic	1	2	3	4	5
Inspired	1	2	3	4	5
Proud	1	2	3	4	5
Determined	1	2	3	4	5
Strong	1	2	3	4	5
Active	1	2	3	4	5
Upset	1	2	3	4	5
Distressed	1	2	3	4	5
Nervous	1	2	3	4	5
Jittery	1	2	3	4	5
Guilty	1	2	3	4	5
Ashamed	1	2	3	4	5
Hostile	1	2	3	4	5
Irritable	1	2	3	4	5
Scared	1	2	3	4	5
Afraid	1	2	3	4	5

Figure 7: PANAS Scale

PANAVA-KS Scale

Much like the PANAS Scale, the PANAVA-KS Scale, developed by Urs Schallberger in 2005, is based on the PANA Model. However, unlike the PANAS Scale, it introduces a further element, valence, uses bipolar scales and consists of ten items.

How do you feel at the moment?
(Put a cross in each row at the point which best corresponds with your feelings)

	(3)	(2)	(1)	(0)	(1)	(2)	(3)	
	very			undecided			very	
satisfied	O	O	O	O	O	O	O	dissatisfied
full of energy	O	O	O	O	O	O	O	no energy
stressed	O	O	O	O	O	O	O	relaxed
tired	O	O	O	O	O	O	O	wide awake
peaceful	O	O	O	O	O	O	O	angry
unhappy	O	O	O	O	O	O	O	happy
listless	O	O	O	O	O	O	O	highly motivated
calm	O	O	O	O	O	O	O	nervous
enthusiastic	O	O	O	O	O	O	O	bored
worried	O	O	O	O	O	O	O	free of worry

Figure 8: PANAVA-KS Scale

A benefit of the PANAVA-KS Scale is that it employs accessible language suitable for Toolkit users. Due to the format of the scale, it needs minimal adaptation. For example, 'satisfaction' could be used as a sentiment dimension alongside the dimension statement 'It made me feel satisfied'; using a sliding scale the user could select a point on the slider.

A limitation of the PANAVA-KS is that, much like the PANAS Scale, the number of items within the scale would need to be reduced for the Toolkit. In addition, PANAVA-KS is generally used as an Experience Sampling Method (ESM) to measure emotion over time from the same group or individual, rather than comparing the results of multiple individuals for the same event/experience. However, it does present a starting point for measuring sentiment that could be adapted for the Toolkit.

LE-PANAVA Scale

The Lebender Emoticon PANAVA Scale (LE-PANAVA), developed by Marc Schreiber and Gregor Jenny in 2020, is based on the ten item verbal PANAVA-KS Scale. The LE-PANAVA Scale consists of five items capturing positive emotion, negative emotion and valence. The LE-PANAVA Scale has been designed for use on digital devices.
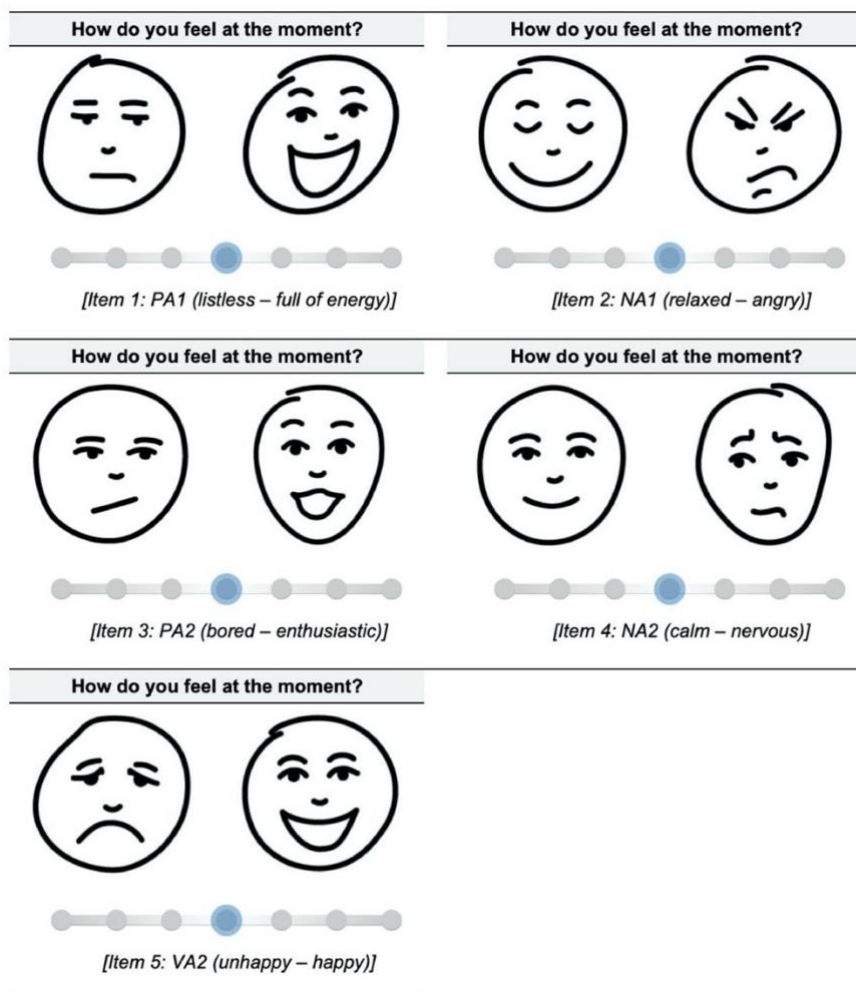

Figure 10: LE-PANAVA Scale

A benefit of the LE-PANAVA Scale is that it presents a concise five item subset of PANAVA-KS and PANAS Scale items, minimising the risk of Toolkit users experiencing survey fatigue. The LE-PANAVA Scale employs accessible language and images suitable for Toolkit users and, much like the PANAVA-KS Scale, due to its format, needs minimal adaptation. A limitation is that the gathered data, as with PANAVA-KS and PANAS Scale, is used to create a final score that needs to be 'read' by Toolkit users to understand whether it is indicating high positive affectivity, low positive affectivity, low negative affectivity or high negative affectivity. However, due to the accessible language used, it has the benefit of individual item metrics being readily understood by Toolkit users.

Comparison of Options

The report will now compare the identified options against the evaluation criteria.

In terms of identifying a complete framework that encompasses positive and negative emotions, Plutchik's Model, in particular the development of the Wheel of Emotion, presents the most nuanced option. Because Plutchik's model is a discrete and dimensional model hybrid, it has the benefit of both a visual interrelated framework and distinct emotional categories. In contrast, the Circumplex Model and PAD Model, despite presenting a complete framework, both require a further step for results to be 'read' and understood. As such, the results may be too abstract for Toolkit users and consequently difficult to interpret. The PANA Model is the only framework that recognises the experience of positive and negative emotions simultaneously, however it is designed to be understood through the generation of a final emotion score. Despite this, due to the accessible language used in the PANA Model's associated affect measures, the PANAVA-KS Scale and LE-PANAVA Scale, Toolkit users would be more readily able to understand individual metric results as well as derive more complex emotional analysis from the final emotion score.

In terms of identifying a framework suitable to be adapted to dimension statements within the Toolkit, Plutchik's Model would need to be radically adapted, as it does not have a closely associated affect measure. For example, you could identify bipolar items such as 'joy' and 'sadness' and present them on a sliding scale. However, due to the nature of the framework, not all emotions have a clear opposite within the wheel, for example 'fear' is opposite 'anger'. In contrast, the Circumplex Model and PAD Model both present a clear way to measure a full range of emotions through two or three dimension questions respectively, however the language used would need to be adapted for Toolkit users. AS and SAM presents a non-verbal way that Circumplex Model and PAD Model language could be made more accessible. The PANA Model, despite being the basis for the widely used PANAS Scale affect measure, doesn't initially present itself as easily adaptable to Toolkit use due to the language used and its twenty individual items. However, the PANAVA-KS Scale and LE-PANAVA Scale, consisting of ten and five items respectively, present accessible bipolar scales that require minimal adaptation.

Final Recommendation

The four options identified within this report all present benefits and limitations. As such, this report recommends implementing the PANA Model in combination with the LE-PANAVA Scale. The PANA Model presents a complete, readily understandable, sentiment framework that can be measured using only five items in the LE-PANAVA Scale.

The LE-PANAVA Scale can be adapted to the Toolkit in several ways. For example:

- Each of the five items could be adapted into separate dimensions.
- The LE-PANAVA Scale could be adapted to a dimension statement and the five items contained within as a question matrix (e.g. 'Sentiment: It made me feel a particular way').
- The LE-PANAVA Scale could remain unadapted and be added to the platform using the terminology 'Affect Measure'.

If the LE-PANAVA Scale is adapted into five separate dimensions, guidance would need to be created regarding the importance of using all five dimensions to make full use of the PANA framework. Dimension questions, on average, take a respondent five seconds to answer, meaning the scale would take respondents twenty-five seconds to answer in total.

A limitation of all affect measures is that they are designed to capture the current emotion of a respondent. This needs to be considered if respondents are completing a survey a long time after an event. Wording may need to be used to indicate how the respondent should answer, for example, rather than 'How do you feel at the moment?' being used, the past tense 'How did you feel during the event?' could be used.

Conclusion

To conclude, this report recommends implementing the LE-PANAVA Scale. The LE-PANAVA Scale has a concise framework, captures both positive and negative affect and presents several options for Toolkit dimension adaptation.

Going forward CWC will need to consider not only how the LE-PANAVA Scale is adapted as a Toolkit dimension, but also how the subsequent collected data will be presented to Toolkit users on the Culture Counts platform. Decisions around how the data is accessed by users, for example, in the form of charts or as raw data, will impact what support guidance needs to be developed.

Support guidance developed will, in general, need to consider: the relationship between the affect measure, the LE-PANAVA Scale, and the emotion classification, the PANA Model; the importance of using the full five item LE-PANAVA Scale; how to gain insights from the collected data.

References

Bales, R. F., (2001) *Social Interaction Systems.* United States: Transaction Publishers. Available at https://www.google.co.uk/books/edition/Social_Interaction_Systems/gQWsAAAAQBAJ?hl=en&gbpv=0 (Accessed: 07 June 2021)

Bastian, B., Harmon-Jones, C., and Harmon-Jones, E., (2016) 'The Discrete Emotions Questionnaire: A New Tool for Measuring State Self-Reported Emotions'*, PLoS ONE.* Available at: https://journals.plos.org/plosone/article/comment?id=10.1371/annotation/b16ea8c3-dd8b-4421-b793-e2752f09926f (Accessed: 22 June 2021)

Biswas-Diener, R., Choi, D., Diener, E., Kim-Prieto, C., Oishi, S., Tov, W., and Wirtz, D., *(*2010*)* 'New well-being measures: Short scales to assess flourishing and positive and negative feelings', *Social Indicators Research*. Available at: https://link.springer.com/article/10.1007%2Fs11205-009-9493-y (Accessed: 23 June 2021)

Bradley, M. M., and Lang, P. J., (1994) 'Measuring Emotion: The Self-Assessment Manikin and The Semantic Differential', *Journal of Behaviour Therapy and Experimental Psychiatry*. Available at https://edwp.educ.msu.edu/wp-content/uploads/sites/29/2020/06/Bradley_1994_MeasuringEmotionTheSelfassessment.pdf (Accessed: 22 June 2021)

Cerin, E., Szabo, A., and Williams, C., (2001) 'Is the Experience Sampling Method (EMS) Appropriate for Studying Pre-Competitive Emotions?', *Psychology of Sport and Exercise Vol 2 Issue 1*. Available at https://www.sciencedirect.com/science/article/abs/pii/S1469029200000091 (Accessed: 22 June 2021)

Chatterjee, H. J., and Thomson, L. J., (2013) *UCL Museum Wellbeing Measures Toolkit.* Available at https://www.ucl.ac.uk/culture/sites/culture/files/ucl_museum_wellbeing_measures_toolkit_sept2013.pdf (Accessed 23 August 2021)

Choi, Y., Skiena, S., Ward, C., and Xavier, E. (2011) 'Empath: A Framework for Evaluating Entity-Level Sentiment Analysis', *8th International Conference & Expo on Emerging Technologies for a Smarter World.* Available at https://ieeexplore.ieee.org/abstract/document/6135866 (Accessed: 07 June 2021)

Clark L. A., Tellegen, A., and Watson, D., (1988) 'Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales', *Journal of Personality and Social Psychology Vol 54.* Available at https://psycnet.apa.org/buy/1988-31508-001 (Accessed: 23 June 2021)

Cowen, A. S., and Keltner, D., (2017) 'Self-Report Captures 27 Distinct Categories of Emotion Bridged by Continuous Gradients', *Proceedings of the National Academy of Sciences.* Available at https://www.pnas.org/content/114/38/E7900 (Accessed: 14 June 2021)

Gregor, J. J., and Schreiber, M., (2020) 'Development and Validation of the 'Lebender Emoticon PANAVA' Scale (LE-PANAVA) for Digitally Measuring Positive and Negative Activation, and Valence via Emoticons', *Personality and Individual Differences,* Vol 160. Available at https://www.sciencedirect.com/science/article/pii/S0191886920301124 (Accessed: 14 June 2021)

Harley, J. M., (2016) 'Measuring Emotions: A Survey of Cutting Edge Methodologies Used in Computer-Based Learning Environment Research', *Emotions, Technology, Design, and Learning.* Available at https://www.sciencedirect.com/science/article/pii/B9780128018569000050 (Accessed: 22 June 2021)

Kazmaier, J., and Vuuren, J. H., (2020) 'A Generic Framework for Sentiment Analysis: Leveraging Opinion-Bearing Data to Inform Decision Making'*, Decision Support Systems Vol 135*. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0167923620300592 (Accessed: 22 June 2021)

Lee, L., and Pang, B., (2008) 'Opinion Mining and Sentiment Analysis', *Foundations and Trends in Information Retrieval Vol. 2*. Available at http://www.cs.cornell.edu/home/llee/omsa/omsa.pdf (Accessed: 07 June 2021)

Liu, B., and Zhang, L., (2012) *A Survey of Opinion Mining and Sentiment Analysis*. Chicago: University of Illinois. Available at http://www.cs.unibo.it/~montesi/CBD/Articoli/SurveyOpinionMining.pdf (Accessed: 07 June 2021)

Mehrabian, A., (1980). *Basic Dimensions for a General Psychological Theory*. Cambridge: Gunn & Hain. Available at https://archive.org/details/basicdimensionsf0000mehr/page/39/mode/2up (Accessed: 07 June 2021)

Plutchik, R., (1991) *The Emotions.* United States: University Press of America. Available at https://www.google.co.uk/books/edition/The_Emotions/JaQauznPoiEC?hl=en&gbpv=0 (Accessed: 07 June 2021)

Rubin, D., and Talarico, J., (2009) 'A Comparison of Dimensional Models of Emotion: Evidence from Emotions, Prototypical Events, Autobiographical Memories, and Words', *Memory*. Available at https://www.tandfonline.com/doi/abs/10.1080/09658210903130764 (Accessed: 07 June 2021)

Russell, J. A., (1980) 'A Circumplex Model of Affect', *Journal of Personality and Social Psychology*. Available at https://doi.org/10.1037/h0077714 (Accessed: 07 June 2021)

Russell, J. A., and Barrett, L. F., (1999) 'Core Affect, Prototypical Emotional Episodes, and Other Things Called Emotion: Dissecting the Elephant', *Journal of Personality and Social Psychology*. Available at https://doi.org/10.1037/0022-3514.76.5.805 (Accessed: 07 June 2021)

Turner, J., (2000) *On the Origins of Human Emotions: A Sociological Inquiry into the Evolution of Human Affect.* United States: Stanford University Press. Available at https://www.google.co.uk/books/edition/On_the_Origins_of_Human_Emotions/aEeSmDRsXkcC?hl=en&gbpv=0 (Accessed: 07 June 2021)

Watson, D., & Tellegen, A., (1985) 'Toward a Consensual Structure of Mood', *Psychological Bulletin*. Available at https://doi.org/10.1037/0033-2909.98.2.219 (Accessed: 07 June 2021)

Zhang, W., and Fesenmaier, D. R., (2018) 'Assessing Emotions in Online Stories: Comparing Self-report and Text-based Approaches', *Information Technology & Tourism*. Available at: https://www.researchgate.net/publication/327705342_Assessing_emotions_in_online_stories_comparing_self-report_and_text-based_approaches (Accessed: 22 June 2021)