Incorporating Emotions into Consumer Evaluation Models:

Application to Kmart Australia

List of web appendices

Web Appendix 1:	Background on discount department stores in Australia
Web Appendix 2:	Industry practice in the measurement of emotions
Web Appendix 3:	Summary of procedure used to development dynamic avatars
Web Appendix 4:	Operationalization and descriptive statistics of key constructs (Refer to the separate Excel spreadsheet)
Web Appendix 5:	Principal Components Factor Analysis of Emotional Responses
Web Appendix 6:	Structure of workshops to translate research insights into action
Web Appendix 7:	Details of the Aggregate Store Visits Model to Test Effect of Copy Design
Web Appendix 8:	Applications of Methodology to Measure Emotions and their Effect of Choice

Web Appendix 1: Discount Department Stores in Australia

The keen competition from omni-channel retailing in Australia has added to the difficult trading terms that have continued beyond the global financial crisis. Over the five years from 2008-09 to 2012-13, discount department store revenue in Australia expanded at a compound annual rate of just 2.1%[1]. Constraining growth has been anaemic consumer spending, fierce price competition and direct sourcing allowing for price deflation. Meanwhile, online retail sales have motored ahead at +14%[2] annual growth, with 45% of the sales substituting for products sold in department stores (including fashion)[3].

In all corners of the sector there is evidence of an industry in transformation. Most evident is consumers searching and shopping online, retailers reducing inventory as they exit unprofitable categories, and price deflation as retailers move their logistics models to embrace direct sourcing.

In 2013 the sector employed 47,230 Australians.[4] Like many sectors of the Australia economy, the discount department store sector is characterized by a high concentration ratio. Indeed, of the estimated $ 9b in sales[5], 95%[6] is represented by Big W and Kmart.

The 363 stores (178 Big W and 185 Kmart) are distributed throughout the six states and three mainland territories of the Australia. Adding to the other pressures is the relative cost of rent. In Australia there is only about half the retail space per capita as in the USA (2.12 sq.m. vs 4.21 sq.m.)[7].

[1] IBISWorld Research Report, Discount Department Stores in Australia: Market Research Report, March 2013, http://www.ibisworld.com.au/industry/discount-department-stores.html

[2] NAB Online Retail Sales Index monthly update – June 2013, 01 August 2013, http://business.nab.com.au/online-retail-sales-index-monthly-update-june-2013-4324/

[3] NAB Online Retail Sales Index: Indepth & Special report – October 2013, 02 December 2013, http://business.nab.com.au/online-retail-sales-index-indepth-special-report-october-2013-5126/

[4] IBISWorld Research Report, Discount Department Stores in Australia: Market Research Report, March 2013, http://www.ibisworld.com.au/industry/discount-department-stores.html

[5] IBISWorld Research Report, Discount Department Stores in Australia: Market Research Report, March 2013, http://www.ibisworld.com.au/industry/discount-department-stores.html

[6] Sales based on company annual reports. IBIS World Industry Report OD5077 "Discount Department Stores in Australia" March 2013 does not classify Target as a Discount Department store, reflecting its mid-tier status.

[7] Shopping Centre Comparison, Shopping Council of Australia, May 2009, http://www.scca.org.au/Pdf%20links/2009PDFlinks/Michael%20Baker.%20Aust-US%20Centre%20Performance%20Comparison%20%20May%2009.pdf

Web Appendix 2: Survey of Industry Applications of Measurement of Emotions

In many areas of marketing science, industry practice leads academic research in terms of sophistication and methodology development (Winer 2015). Rossiter and Percy (2013) suggest that there is a "notable worsening of the academic-practitioner divide" in one important aspect of the research field addressed by this article; advertising design and evaluation. For that reason we attempted to benchmark our research not only against academic, peer-reviewed work in the measurement of emotions, but also against industry practice.

It is difficult to determine the extent, nature and quality of commercial methodologies to gauge emotions for a number of reasons. Firstly, such methodologies are often proprietary and their owners may have a strong interest in preserving their confidentiality. Secondly, few industry studies have been subject to peer review, and thus it is difficult to evaluate the quality and rigor of the research. For example, the websites of a number of the services talk about methods "currently under development" or "in testing stages." Thirdly, because of the nature of trade publications, it is frequently impossible to see the details of the methodologies that such studies use. To address this problem we used a number of approaches. First, we identified the fifty largest global market research companies by revenue and examined a description of services they offer. Second, we used a search engine (Google) to identify commercial services that might specialize in this area by looking for "Market Research Emotions" in an effort to identify firms offering services measuring emotions. Finally, we undertook a search of industry publications to uncover technical advances in the area of calibrating and modeling emotions by industry.

Search of top fifty market research firms self-stated service portfolios

The American Marketing Association annually ranks the top fifty market research firms by revenue and includes a description of the services that they provide (Gold 2014). We searched this list of services for references to "emotions" and "feelings." Six firms' services contain references to the calibration of "emotions," as outlined in Table W2.1, while none refer to the measurement of "feelings." We then visited the websites of these six firms and examined their services, both manually and using internal search engines. In Table W2.1 we note each firm's ranking in the industry (taken from Gold 2014), the name of any proprietary tool where we could identify it, and the measurement approach that the firm appears to use.

Google search of commercial services to calibrate emotions

An alternative way to identify industry practice used to calibrate emotions is to start where a naïve client seeking such services might begin; by entering "market research measuring emotions" into Google search. We analyzed the first three pages of results from this process. While many of the results from this search were to conference papers unrelated to applications, journal articles, and blogs; this approach did identify a number of additional organizations offering services to measure emotions. The companies and their associated tools (taken from a

drill down through the url identified by Google) from the first page of the search (in order of appearance) are presented in Table W2.2.

Table W2.1 Tools identified by a search of industry top research companies (Gold 2014)

Rank	***Company***	***Proprietary tool***	***Approach***
20	Lieberman Research	Sensory Virtual Research	Use of avatars to represent emotions
22	Communispace	Emotion Centric Explorer	Direct elicitation of emotions
29	Perception Research Services	PRS Eye Tracking	Eye tracking and facial coding
36	Radius Global Market Research		Direct elicitation of emotions (Psychographics and semiotics)
43	Gongos Inc		Metaphors and story telling
46	Pert Group	Neuro GPS	Neurological measurement and direct elicitation

Table W2.2 Tools identified by search of "Market research measuring emotions" in Google

Order appearing	***Company***	***Proprietary tool***	***Approach***
1	Affectiva	Affdex	Galvonic skin response (GSR) and facial coding
2	QRI Consulting	Quali-Quant	Projective research
3	GfK	EMO Sensor	Facial coding
4	RealityMine (allied with Kantor, Millward Brown and Affectiva)		Facial coding
5	Research-Live	Emotional Signature	Projective (Drawing, role plays and storytelling); direct elicitation
6	nviso	Emotion Video Analytics	Facial coding
7	Ipsos-Mori	Neuro GPS	Heart rate, GSR, Respiratory and heart rate, Motion, Eye tracking

Industry publications relating to emotions

Rossiter and Percy (2013) note 'the glaring lack of citations of trade publication of trade publications such as *Advertising Age* and *AdMap* in journal articles and textbooks" so we undertook a Google Scholar search of "Advertising Age Emotions Measure" and "AdMap Emotions Measure" and examined the first three pages of results for each. The *Journal of*

Advertising and the *Journal of Advertising Research* are both journals with a substantial industry readership and author base, as well as being relevant to the topic of this research, so we also searched "Journal of Advertising (Research) Emotions", using Google. Finally, ESOMAR is a worldwide market research professional body, consisting primarily of industry practitioners, with 4900 members spread across 130 countries. We reviewed the Esomar Research Paper Series, also using Google Scholar.

Unsurprisingly, each of these sources revealed large number of articles on the measurement of emotions. There were many excellent reviews articles (e.g., Cooper and Pawle 2005, Micu and Plummer 2010). Also academic articles to illustrate new methodologies were also in evidence (e.g., Rossiter and Bellman 2012). Finally, each of the major approaches to measuring emotions are described: direct elicitation (Malhotra *et al*. 2012), physiological methods (Gordon *et al.* 2013), neurological methods (Stipp 2014), and projective methods (Gregg *et al.* 2013). Of particular interest to this research is the use of metaphors including emoticons and avatars. Examples of this work include Morris (1995), Roberts (2013) and Penn and Lughart (2014). Unfortunately, in most cases the actual scales used and their validation and testing are not described, but what the survey does reveal, is that the measurement of emotions occupies a major role in commercial market research. The vast majority of this research is focused on emotions as an end in themselves, addressing questions such as "Does our brand/advertisement elicit feelings of (say) pride?" The use of emotions as an input to models of evaluation and choice in industry seems to be very sparsely populated indeed.

Summary of Industry Applications

An examination of Table W2.1 reveals that all three of the major methods of calibrating emotions identified in Section 3.2 are being used in industry. This diversity of approaches is typical of an industry early in its product life cycle (till a dominant paradigm emerges).

We understand that our approach may under-represent the penetration of emotion measurement by industry for the reasons given at the beginning of this Web Appendix. We do not pretend that this list is complete, only that it was generated using explicit, reasonable, and objective criteria. For example, an examination of the website of the world's largest market research company, Nielsen Holdings NV, shows that it has a service called NeuroFocus that uses facial recognition and EEG, (even though that service is not identified by any of the searches undertaken above).

However, we do believe that the approach does give a flavor of the range of approaches being adopted by industry. It is notable that, by and large, the firms nominating emotion measurement as one of their tools are not the biggest in the industry. (The largest in Table W2.1 is Lieberman Research at number 20.) This would appear to suggest that the largest market research firms do not yet see emotion measurement as one of their key selling points to their clients and potential clients. It is also interesting that in most applications the objective of these methodologies appears to be to identification of emotions elicited by marketing stimuli such as brand names and

communications. That is, the calibration of emotions is an end in itself. That role is in contrast to our objective which is to use emotions as an antecedent to models of future consumer behavior. While it may exist, we could find no evidence of emotions elicited from consumers being incorporated into models of past behavior or future behavioral intentions in the spirit of marketing science choice models.

References

Berger, J., & Milkman, K. L. (2012). "What makes online content viral?" *Journal of Marketing Research*, *49*(2), 192-205.

Cooper, Peter, and John Pawle (2005) "Measuring Emotion in Brand Communication" Paris: ESOMAR Innovate Conference, pp. 1-25

Gold, Laurence N. (2014) "AMA Gold Top 50 Report" Marketing *News* (June) pp. 28-104

Gordon, Alastair, Joe Wheller, Vuong Manh Giang and Winnie Yeung (2013) "Tuning into TV ads: How emotion research transforms our understanding of advertising in Asia," ESOMAR, Asia Pacific, Ho Chi Minh City, April

Gregg, A. P., J. Klymowsky, D. Owens, and A. Perryman (2013). "Let their fingers do the talking? Using the Implicit Association Test in market research" *International Journal of Market Research*, *55*(4), 487-503.

Malhotra, Naresh, Soumya Mukhopadhyay, Xiaoyan Liu and Satyabhusan Dash (2012) "One, few or many? An integrated framework for identifying the items in measurement scales" *International Journal of Market Research*, Vol. 54, No. 6, pp. 835-862

Micu, Anca Cristina and Joseph T. Plummer (2010) "Measurable Emotions: How Television Ads Really Work - Patterns of Reactions to Commercials Can Demonstrate Advertising Effectiveness" *Journal of Advertising Research*, Vol. 50, No. 2, pp. 137-153

Morris, Jon (1995) "Observations: SAM: The Self-Assessment Manikin - An Efficient Cross-Cultural Measurement of Emotional Response" *Journal of Advertising Research*, Vol. 35, No. 6, November/December

Penn, David and Suzanne Lugthart (2014) "Our lips are sealed: Why the Truth s Implicit" Esomar Congress Proceedings pp. 4-24

Roberts, Ken (2013) "A revolution in ad testing" *Admap* July-August pp 10-12

Rossiter, John and Steven Bellman. (2012) "Emotional branding pays off: How brands meet share of requirements through bonding, companionship, and love" *Journal of Advertising Research* (September) 291-296

Rossiter, John R., and Larry Percy (2013) "Observations: How the roles of advertising merely appear to have changed." *International Journal of Advertising* 32.3 pp 391-398.

Stipp, Horst (2014) "Applied neuroscience: From novelty to must-use" *Admap,* September, pp. 38-40

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-induced engagement in internet video advertisements. *Journal of Marketing Research*, *49*(2), 144-159.

Winer, Russell S. (2014). "The Impact of Marketing Science Research on Practice: Comment," *International Journal of Research in Marketing*, 31.2 (2014): 142-143.

Web Appendix 3: Procedure to develop dynamic avatars used in the research

To develop avatars which represented emotions felt by respondents in relation to marketing stimuli, we first ran five focus groups in which respondents were asked to describe the image that came to their minds in relation to a list of emotions, including the nine identified emotions used in this research. From a total of approximately 50 participants, a list of approximately 250 images was generated. We categorized these according to the metaphorical domains associated with the images. We found that similar underlying metaphorical themes were expressed across the images described by a range of different individuals for each targeted emotion. For example, the emotion of happiness was very commonly associated with "being up", as well as with "warm and sunny" images. These observations led to the selection of a final list of three candidate images per emotion.

We then tested three candidate images for each emotion to select a single final image for each emotion. The initial pilot test of these images involved 31 participants, with the goals of finding the most effective design format and identifying difficulties encountered in completing the required tasks, including ease of understanding of the concepts behind the candidate images. The test methodology was based on the Implicit Association Test (IAT) (Greenwald, McGhee, and Schwartz 1998). We programmed the candidate images using the Inquisit™ software package from Millisecond Software™ (www.millisecond.com).

The three candidates for each emotion were also tested in a further study, in which the objective was to use a reaction time task (RTT) or response latency to identify which of the three images would be most strongly associated with the targeted emotion (Bluemke and Friese 2008). Each respondent was required to complete both the RTT and a series of self-report questions. The study was conducted on line using a panel of 750 respondents, recruited by a digital data collection agency, GMI (www.gmi-mr.com). Each panel member was sent a random selection of five Inquisit™ test links, with each RTT having a quota of 150 respondents out of the total panel of 750. The data gathered during the study were validated and cleaned according to procedures set out for the IAT. In particular, trials with a reaction time of greater than 10 seconds were deleted.

The results were processed via an analysis of variance (ANOVA) model. For emotions where no avatar provided a strong and unambiguous representation, a combination of visual concepts drawn from groups with the lowest response latency and highest self-report scores were

used. Animated visual scales were then developed by which respondents could change the animations using a sliding scale operated with a computer mouse to reflect their feelings towards a given stimulus. We used Adobe™ Flash™/Shockwave™ to facilitate web-based delivery of the interactive visual scales. This resulted in 11 pictures for each avatar ranging from the neutral position all the way to extreme emotion.

Reference

Bluemke, Matthias and Malte Friese (2008) "Reliability and validity of the Single-Target IAT (ST-IAT): assessing automatic affect towards multiple attitude objects." *European Journal of Social Psychology* 38 6: 977-997.

Web Appendix 4: Variable definitions and marginal distribution

Jun-11

		Kmart						Target						Big W						Skewness		Kurtosis			
		Mean	Standard	Valid N	Median	Minimum	Maximu	Mean	Standard	Valid N	Median	Minimum	Maximum	Mean	Standard	Valid N	Median	Minimum	Maximum	Statistic	Std. Error	Statistic	Std. Error	Description of measure	Scale Anchors
Feelings	Surprise	0.66	3.97	329	0	-10	10	0.44	3.83	297	0	-10	10	0.48	3.93	326	0	-10	10	-0.06	0.08	0.35	0.16	See url in text of paper	No text anchors
	Happiness	0.06	3.72	329	0	-10	10	-0.73	3.78	297	0	-10	10	0.07	3.81	326	0	-10	10	-0.04	0.08	0.26	0.16		No text anchors
	Love	-0.30	3.91	329	0	-10	10	-0.68	4.11	297	0	-10	10	0.19	3.67	326	0	-10	10	-0.04	0.08	0.41	0.16		No text anchors
	Pride	0.06	3.70	329	0	-10	10	-0.45	3.65	297	0	-10	9	0.01	3.67	326	0	-10	10	-0.02	0.08	0.29	0.16		No text anchors
	Contentment	-0.38	3.89	329	0	-10	10	-0.85	3.97	297	-1	-10	10	-0.40	3.95	326	0	-10	10	-0.02	0.08	0.06	0.16		No text anchors
	Anger	-0.30	2.68	329	0	-10	10	-0.06	2.41	297	0	-10	10	-0.32	2.52	326	0	-10	9	-0.78	0.08	5.78	0.16		No text anchors
	Sadness	-0.70	3.01	329	0	-10	10	-0.22	2.78	297	0	-10	10	-0.80	2.81	326	0	-10	7	-0.89	0.08	3.37	0.16		No text anchors
	Anxiety	-0.59	3.10	329	0	-10	10	-0.33	2.89	297	0	-10	10	-0.68	2.90	326	0	-10	10	-0.56	0.08	2.84	0.16		No text anchors
	Shame	-0.47	3.07	329	0	-10	10	-0.33	2.80	297	0	-10	10	-0.90	2.81	326	0	-10	7	-0.62	0.08	2.69	0.16		No text anchors
	Value	7.24	1.97	329	7.4	0	10	7.29	1.84	297	7.5	0.1	10	7.42	1.87	326	7.7	0	10	-0.89	0.08	1.14	0.16	0 to 10 scale	Poor value, Excellent value
	Quality	7.47	1.77	329	7.8	0	10	7.58	1.74	297	7.9	0.1	10	7.60	1.81	326	7.9	0	10	-0.99	0.08	1.62	0.16	0 to 10 scale	Poor quality, Excellent quality
	Price	7.47	1.78	329	7.8	0.2	10	7.09	1.95	297	7.5	0	10	7.58	1.78	326	7.8	0.1	10	-0.90	0.08	1.01	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
	Performance	7.38	1.98	329	7.7	0.1	10	7.38	1.84	297	7.8	0.1	10	7.54	1.89	326	7.8	0	10	-1.05	0.08	1.42	0.16	0 to 10 scale	Poor performance, Excellent performance
	Reputation	7.69	1.81	329	8.1	0	10	7.72	1.85	297	8.1	0.1	10	7.79	1.85	326	8.15	0	10	-1.21	0.08	2.01	0.16	0 to 10 scale	Poor reputation, Excellent reputation
Performance Model	Available help	5.60	2.52	329	5.6	0	10	5.68	2.65	297	6	0	10	5.82	2.51	326	5.8	0	10	-0.36	0.08	-0.49	0.16	0 to 10 scale	Poor performance, Excellent performance
	Cheerful staff	6.55	2.31	329	6.9	0	10	6.57	2.38	297	7	0.1	10	6.68	2.34	326	6.8	0	10	-0.70	0.08	0.19	0.16	0 to 10 scale	Poor performance, Excellent performance
	Low waiting times	6.20	2.34	329	6.3	0	10	5.95	2.51	297	6.1	0	10	6.28	2.39	326	6.5	0	10	-0.57	0.08	-0.15	0.16	0 to 10 scale	Poor performance, Excellent performance
	A store I feel comfortable in	7.58	1.98	329	7.9	0	10	7.57	2.01	297	8	0	10	7.83	1.89	326	8.1	0.1	10	-1.17	0.08	1.62	0.16	0 to 10 scale	Poor performance, Excellent performance
	Products are on the shelves	6.85	2.13	329	7	0	10	6.99	2.05	297	7.4	0	10	7.13	1.92	326	7.4	0	10	-0.90	0.08	0.90	0.16	0 to 10 scale	Poor performance, Excellent performance
	Good shelf layouts	7.09	2.11	329	7.4	0	10	7.04	2.17	297	7.4	0	10	7.39	1.87	326	7.6	0	10	-0.93	0.08	0.84	0.16	0 to 10 scale	Poor performance, Excellent performance
	Products that last	6.80	2.02	329	7	0	10	7.23	1.99	297	7.5	0.1	10	7.00	2.02	326	7.4	0	10	-0.72	0.08	0.43	0.16	0 to 10 scale	Poor performance, Excellent performance
	Aesthetically pleasing products	7.18	1.83	329	7.3	0	10	7.63	1.71	297	8	0.1	10	7.32	1.85	326	7.6	0	10	-0.76	0.08	0.84	0.16	0 to 10 scale	Poor performance, Excellent performance
	Good location	7.10	2.59	329	7.7	0	10	7.42	2.24	297	7.9	0	10	6.91	2.68	326	7.55	0	10	-1.02	0.08	0.46	0.16	0 to 10 scale	Poor performance, Excellent performance
	Excellent opening hours	8.15	1.59	329	8.5	1.6	10	7.67	1.78	297	8.1	2.3	10	7.86	1.61	326	8.1	0	10	-0.79	0.08	0.30	0.16	0 to 10 scale	Poor performance, Excellent performance
	Accepts products back	7.57	1.75	329	8	2.2	10	7.54	1.85	297	7.9	0.1	10	7.58	1.87	326	7.9	0	10	-0.76	0.08	0.38	0.16	0 to 10 scale	Poor performance, Excellent performance
	Rewards me as a customer	6.05	2.50	329	6.1	0	10	5.85	2.45	297	5.7	0	10	6.13	2.41	326	5.9	0	10	-0.41	0.08	-0.23	0.16	0 to 10 scale	Poor performance, Excellent performance
Reputation Model	Many types of products	7.50	1.92	329	7.8	0	10	7.46	1.87	297	7.8	0.1	10	7.74	1.77	326	8	0.1	10	-1.07	0.08	1.59	0.16	0 to 10 scale	Poor reputation, Excellent reputation
	Good brands	6.76	2.02	329	7	0	10	7.04	2.04	297	7.2	0	10	7.02	1.99	326	7.25	0	10	-0.66	0.08	0.36	0.16	0 to 10 scale	Poor reputation, Excellent reputation
	A store for people like me	7.52	2.05	329	7.9	0	10	7.48	2.03	297	7.8	0	10	7.68	2.01	326	8.1	0	10	-1.07	0.08	1.23	0.16	0 to 10 scale	Poor reputation, Excellent reputation
	Family-oriented	7.81	1.74	329	8.2	0.4	10	7.69	1.91	297	8.1	0	10	7.97	1.68	326	8.3	0.1	10	-1.17	0.08	1.98	0.16	0 to 10 scale	Poor reputation, Excellent reputation
	A store I believe in	7.30	1.89	329	7.6	0.1	10	7.55	1.89	297	7.9	0	10	7.47	1.90	326	7.8	0	10	-0.89	0.08	1.03	0.16	0 to 10 scale	Poor reputation, Excellent reputation
	Fun to look around	7.59	1.99	329	8.1	0	10	7.65	1.94	297	8.2	0	10	7.86	1.83	326	8.3	0.1	10	-1.14	0.08	1.45	0.16	0 to 10 scale	Poor reputation, Excellent reputation
Price Model	Good promotions	7.48	1.91	329	8	0.1	10	7.44	1.90	297	7.8	0	10	7.44	1.86	326	7.75	0	10	-0.93	0.08	1.03	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
	Low catalogue prices	7.23	1.96	329	7.5	0.1	10	7.34	1.81	297	7.6	0	10	7.26	1.95	326	7.6	0	10	-0.78	0.08	0.74	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
	Easy to understand prices	7.37	1.79	329	7.5	0.1	10	7.18	1.94	297	7.5	0.1	10	7.44	1.83	326	7.7	0	10	-0.72	0.08	0.59	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
	Always low prices	7.49	1.89	329	7.8	0	10	6.87	2.05	297	7.1	0.1	10	7.63	1.95	326	8.1	0.2	10	-0.96	0.08	0.96	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
	Will beat competitors	6.58	2.17	329	6.6	0	10	6.37	2.19	297	6.6	0	10	6.67	2.27	326	6.7	0	10	-0.58	0.08	0.33	0.16	0 to 10 scale	Not at all price competitive, Highly price competitive
Store Choice	Likelihood to choose store	7.64	2.29	329	8.2	0	10	7.51	2.16	297	8.1	0	10	7.65	2.25	326	8.2	0.1	10	-1.17	0.08	1.08	0.16	0 to 10 scale	Extremely Unlikely. Extremely Likely

Jun-12

		Kmart						Target						Big W						Skewness		Kurtosis			
		Mean	Standard	Valid N	Median	Minimum	Maximu	Mean	Standard	Valid N	Median	Minimum	Maximum	Mean	Standard	Valid N	Median	Minimum	Maximum	Statistic	Std. Error	Statistic	Std. Error	Description of measure	Scale Anchors
Feelings	Surprise	0.71	4.10	759	0	-10	10	0.49	3.98	759	0	-10	10	0.54	4.02	759	0	-10	10	0.08	0.05	0.20	0.10	See url in text of paper	No text anchors
	Happiness	0.33	4.03	759	0	-10	10	-0.06	3.98	759	0	-10	10	0.14	4.07	759	0	-10	10	0.09	0.05	0.16	0.10		No text anchors
	Love	-0.04	3.99	759	0	-10	10	-0.35	4.03	759	0	-10	10	-0.23	3.96	759	0	-10	10	-0.02	0.05	0.41	0.10		No text anchors
	Pride	0.11	3.85	759	0	-10	10	-0.23	3.78	759	0	-10	10	-0.06	3.69	759	0	-10	10	0.00	0.05	0.31	0.10		No text anchors
	Contentment	-0.19	4.22	759	0	-10	10	-0.60	3.99	759	0	-10	10	-0.30	4.08	759	0	-10	10	0.07	0.05	0.13	0.10		No text anchors
	Anger	-0.13	2.56	759	0	-10	10	0.12	2.75	759	0	-10	10	-0.17	2.69	759	0	-10	10	-0.27	0.05	4.89	0.10		No text anchors
	Sadness	-0.57	3.07	759	0	-10	10	-0.41	3.03	759	0	-10	10	-0.50	3.04	759	0	-10	10	-0.50	0.05	2.73	0.10		No text anchors
	Anxiety	-0.42	2.91	759	0	-10	10	-0.19	3.02	759	0	-10	10	-0.36	2.87	759	0	-10	10	-0.38	0.05	2.90	0.10		No text anchors
	Shame	-0.36	3.02	759	0	-10	10	-0.14	3.11	759	0	-10	10	-0.34	2.97	759	0	-10	10	-0.40	0.05	2.46	0.10		No text anchors
	Value	7.52	2.16	759	7.93	0	10	7.31	2.09	759	7.67	0	10	7.40	2.05	759	7.72	0	10	-0.86	0.05	0.66	0.10	0 to 10 scale	Poor value, Excellent value
	Quality	7.33	2.18	759	7.64	0	10	7.70	2.03	759	8.22	0	10	7.35	2.08	759	7.65	0	10	-0.92	0.05	0.84	0.10	0 to 10 scale	Poor quality, Excellent quality
	Price	7.98	2.03	759	8.52	0	10	6.91	2.22	759	7.12	0	10	7.65	1.95	759	8	0	10	-1.01	0.05	1.06	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
	Performance	7.50	2.10	759	7.85	0	10	7.35	2.08	759	7.71	0	10	7.41	2.07	759	7.66	0	10	-0.90	0.05	0.89	0.10	0 to 10 scale	Poor performance, Excellent performance
	Reputation	7.58	2.09	759	8.01	0	10	7.73	2.06	759	8.25	0	10	7.65	1.98	759	8	0	10	-0.97	0.05	0.95	0.10	0 to 10 scale	Poor reputation, Excellent reputation
Performance Model	Available help	6.03	2.85	759	6.38	0	10	5.91	2.87	759	6.09	0	10	5.86	2.83	759	6.21	0	10	-0.45	0.05	-0.65	0.10	0 to 10 scale	Poor performance, Excellent performance
	Cheerful staff	6.75	2.60	759	7.06	0	10	6.62	2.59	759	6.92	0	10	6.62	2.54	759	6.9	0	10	-0.70	0.05	-0.06	0.10	0 to 10 scale	Poor performance, Excellent performance
	Low waiting times	6.62	2.56	759	6.85	0	10	6.17	2.64	759	6.5	0	10	6.48	2.52	759	6.72	0	10	-0.57	0.05	-0.28	0.10	0 to 10 scale	Poor performance, Excellent performance
	A store I feel comfortable in	7.76	2.13	759	8.28	0	10	7.51	2.31	759	8.03	0	10	7.61	2.23	759	8.02	0	10	-1.09	0.05	1.06	0.10	0 to 10 scale	Poor performance, Excellent performance
	Products are on the shelves	7.20	2.32	759	7.63	0	10	7.06	2.27	759	7.45	0	10	7.25	2.22	759	7.55	0	10	-0.84	0.05	0.49	0.10	0 to 10 scale	Poor performance, Excellent performance
	Good shelf layouts	7.40	2.17	759	7.72	0	10	7.03	2.31	759	7.32	0	10	7.28	2.23	759	7.66	0	10	-0.82	0.05	0.36	0.10	0 to 10 scale	Poor performance, Excellent performance
	Products that last	6.91	2.31	759	7.08	0	10	7.59	2.12	759	7.96	0	10	7.06	2.18	759	7.34	0	10	-0.74	0.05	0.26	0.10	0 to 10 scale	Poor performance, Excellent performance
	Aesthetically pleasing products	7.31	2.15	759	7.56	0	10	7.78	1.97	759	8.13	0	10	7.29	2.08	759	7.49	0	10	-0.85	0.05	0.64	0.10	0 to 10 scale	Poor performance, Excellent performance
	Good location	7.11	2.82	759	7.91	0	10	7.33	2.48	759	7.95	0	10	6.81	2.86	759	7.34	0	10	-0.96	0.05	0.16	0.10	0 to 10 scale	Poor performance, Excellent performance
	Excellent opening hours	8.18	1.95	759	8.81	0	10	7.63	2.07	759	7.99	0	10	7.66	2.08	759	8	0	10	-0.99	0.05	0.85	0.10	0 to 10 scale	Poor performance, Excellent performance
	Accepts products back	7.76	2.09	759	8.39	0	10	7.61	2.16	759	8.08	0	10	7.54	2.15	759	7.9	0	10	-0.85	0.05	0.40	0.10	0 to 10 scale	Poor performance, Excellent performance
	Rewards me as a customer	6.60	2.54	759	6.68	0	10	6.43	2.64	759	6.58	0	10	6.24	2.62	759	6.18	0	10	-0.53	0.05	-0.25	0.10	0 to 10 scale	Poor performance, Excellent performance
Reputation Model	Many types of products	7.67	2.13	759	8.13	0	10	7.54	2.10	759	7.88	0	10	7.73	2.03	759	8.18	0	10	-0.99	0.05	0.87	0.10	0 to 10 scale	Poor reputation, Excellent reputation
	Good brands	7.06	2.29	759	7.39	0	10	7.16	2.30	759	7.54	0	10	7.25	2.20	759	7.63	0	10	-0.73	0.05	0.13	0.10	0 to 10 scale	Poor reputation, Excellent reputation
	A store for people like me	7.63	2.28	759	8.22	0	10	7.35	2.43	759	7.89	0	10	7.55	2.30	759	8.1	0	10	-1.09	0.05	0.92	0.10	0 to 10 scale	Poor reputation, Excellent reputation
	Family-oriented	8.10	1.92	759	8.55	0	10	7.91	1.97	759	8.37	0	10	8.01	1.95	759	8.45	0	10	-1.13	0.05	1.23	0.10	0 to 10 scale	Poor reputation, Excellent reputation
	A store I believe in	7.37	2.20	759	7.66	0	10	7.82	2.09	759	8.4	0	10	7.43	2.10	759	7.76	0	10	-0.94	0.05	0.78	0.10	0 to 10 scale	Poor reputation, Excellent reputation
	Fun to look around	7.76	2.14	759	8.25	0	10	7.70	2.14	759	8.18	0	10	7.76	2.07	759	8.35	0	10	-1.07	0.05	1.00	0.10	0 to 10 scale	Poor reputation, Excellent reputation
Price Model	Good promotions	7.78	2.08	759	8.27	0	10	7.28	2.15	759	7.47	0	10	7.51	2.07	759	7.81	0	10	-0.94	0.05	0.87	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
	Low catalogue prices	7.45	2.19	759	7.86	0	10	7.36	2.14	759	7.68	0	10	7.34	2.09	759	7.63	0	10	-0.80	0.05	0.48	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
	Easy to understand prices	7.81	2.09	759	8.31	0	10	7.32	2.17	759	7.73	0	10	7.42	2.18	759	7.81	0	10	-0.93	0.05	0.62	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
	Always low prices	8.02	2.01	759	8.59	0	10	6.62	2.19	759	6.76	0	10	7.58	1.94	759	7.81	0	10	-0.85	0.05	0.60	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
	Will beat competitors	7.09	2.28	759	7.28	0	10	6.56	2.28	759	6.68	0	10	6.99	2.19	759	7.08	0	10	-0.57	0.05	0.19	0.10	0 to 10 scale	Not at all price competitive, Highly price competitive
Store Choice	Likelihood to choose store	7.78	2.33	759	8.47	0	10	7.57	2.43	759	8.25	0	10	7.74	2.32	759	8.28	0	10	-1.21	0.05	1.18	0.10	0 to 10 scale	Extremely Unlikely. Extremely Likely

Measurement scale for store visit

Web Appendix 5: Convergent and Discriminant Validity of the Emotion Avatar Measures.

In this appendix we establish which emotions we have measured. We have already pointed out that there is not a perfect correspondence between emotions as felt and the words that we use for them. However, it is still useful to attach words to specific affective reactions (as we have done in selecting our nine emotions). Therefore, we need to determine whether the meaning that we have put on our nine animated avatars corresponds to that which respondents would when asked to identify the emotion they associated with each avatar. To establish convergent and discriminant validity we asked 405 respondents to select one of the nine listed emotions they considered to be closest to that represented by each of the nine emotion avatars (they were also permitted to select options of "other" or "none"). We eliminated respondents who did not move the cursor or rated an avatar as other/none, though the substantive findings are not changed if they are included. This led to an average of 353 observations for each avatar. We undertook a principal components factor analysis with varimax rotation on the resultant 9x9 matrix which enables us to see which emotions load on which avatar. Convergent validity is clearly established. For no avatar is the primary factor loading less than 0.95, demonstrating a strong relationship between the avatar and the underlying emotion. Discriminant validity is also strong; in no case is the factor loading of an avatar more than 0.17 for an emotion with which it was not intended to be associated. This series of tests give us confidence that the basis of our measurement approach is valid and reliable. For a detailed description of the measure development see Roberts, Wong and Stein (2011).

FEELING	FACTOR								
	1	2	3	4	5	6	7	8	9
1 PRIDE	0.02	-0.06	0.03	0.10	**0.97**	0.16	0.13	-0.09	0.04
2 SURPRISE	0.04	-0.05	0.02	0.12	0.16	**0.96**	0.11	-0.09	0.10
3 ANGER	0.06	-0.05	-0.05	0.13	0.13	0.11	**0.96**	-0.07	0.13
4 LOVE	0.07	0.02	0.04	**0.97**	0.10	0.11	0.12	0.03	0.13
5 HAPPY	0.11	0.07	0.17	0.13	0.04	0.10	0.13	0.05	**0.95**
6 SHAME	**0.98**	0.06	0.06	0.07	0.02	0.03	0.05	0.07	0.10
7 SADNESS	0.06	0.07	**0.97**	0.04	0.03	0.02	-0.05	0.15	0.16
8 CONTENTMENT	0.06	**0.98**	0.07	0.02	-0.06	-0.05	-0.05	0.13	0.06
9 ANXIETY	0.07	0.13	0.15	0.03	-0.09	-0.09	-0.07	**0.96**	0.05

Correlation matrices are included in Sheet 2 of Web Appendix 4.

Web Appendix6: Creative Workshop To Design Advertising Copy For Television Commercial

Participants: Advertising agency, Forethought Research, Kmart Marketing team

Pre-workshop	*Workshop background*	*Ideation*	*Draft Content*
Research identifies target emotion(s)	Expert gives background on target emotion(s)	Syndicate ideation on communication vehicles	Talent, imagery, story line, tone
Pre-workshop idea generation	History, brand essence and background	Group evaluation & enhancement	Draft animatics and pre-testing
EXAMPLE			
Target emotion: **Pride**	Making ends meet, Doing a good job, Like other mums	In store with other mums making smart choices	Sharing good deals

Web Appendix 7: Details of the Aggregate Store Visits Model to Test Effect of Copy Design

The basic log-log model of advertising effectiveness in equation (1) of the main paper needs modification to accommodate the phenomena we identify in that section.

To accommodate serial correlation, the disturbance term in equation (1) is modeled by

$$u_t = \rho u_{t-1} + \eta_t , \tag{W1}$$

where $|\rho_i| < 1$ and η_t is an independent random disturbance, distributed normally with mean zero and constant variance.

We allow for possible endogeneity by using instrumental variables (Greene 2003). An added complication in our model is serially correlated errors, but Fair (1970) proposes a method for just this situation, that produces consistent parameter estimates. In our case, Fair's two-step procedure is to regress Adv_t on store visits, advertising and price for the time periods *t*-1 and *t*-2. From this regression model, predicted values $\hat{\mathrm{Adv}}_t$ are obtained and these become instrumental variables that are substituted for Adv_t in equation (1).

The final issue on our list is advertising carryover, whereby advertising from previous time periods might influence sales in the current period (Leone 1995). Hanssens, Parsons and Schultz (2001, pp. 142-152) exhibit several methods for modeling advertising carryover. The method that is most compatible with the multiplicative model in equation (1) is Broadbent's (1979) Adstock model. Adstock is essentially an exponential smoothing of the advertising measure. To be consistent with our use of advertising instruments we need to define Adstock as follows.

$$AS_t = \psi AS_{t-1} + (1-\psi)\hat{\mathrm{Adv}}_t , \tag{W2}$$

where AS_t is the Adstock for Kmart at time *t* and ψ is a smoothing parameter bounded between 0 and 1. We now replace the advertising instruments, $\hat{\mathrm{Adv}}_t$, with the Adstock variable in equation (W2) to incorporate possible advertising carryover effects. Leeflang *et al.* (2000, p. 89) show that allowing for carryover in this way means that β^{Adv} should be interpreted as the long-term advertising effect, while $(1-\psi)\beta^{Adv}$ is the short-term effect.

References

Broadbent, Simon (1979), "One Way TV Advertisements Work", *Journal of the Market Research Society*, 21, 3, 139-166.

Fair, Ray C. (1970), "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors", *Econometrica*, 38, 3, 507-516.

Greene, W. H. (2003), *Econometric Analysis*, 5th ed., Englewood Cliffs, NJ: Prentice-Hall.

Hanssens, Dominique M., Leonard J. Parsons and Randall L. Schultz (2001), *Market Response Models: Econometric and Time-Series Analysis*, 2nd ed., Boston, MA: Kluwer Academic Press.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel and Philippe A. Naert (2000), *Building Models for Marketing Decisions*, Boston, MA: Kluwer Academic Publishers.

Leone, Robert P. (1995), "Generalizing What is Known about Temporal Aggregation and Advertising Carryover," *Marketing Science*, 14, 3 (part 2), G141-G150.

Web Appendix 8: Applications of the Methodology to Measure Emotions and their Effect on Choice

Strategic planners within creative agencies have been fast adopters of the approach because of the granularity of the discrete emotions and the ability to test executions quickly and effectively. The approach has three components, each of which could be adopted in concert or separately by industry. The first element is the use of dynamic avatars. While the specific avatars in this application are subject to a U.S. patent, there are many other static avatars available (e.g., Morris 1995). Second, using the calibration of emotions as an input to structured workshops to design marketing stimuli, making it an antecedent to marketing management rather than a postscript, is clear available to all. Finally, the incorporation of emotions, however measured, into models of choice and behavioral intentions, allows their effect on marketing outcomes to be assessed.

Category	Client companies	Category	Client companies
Packaged goods		***Services***	
Confectionary	Hershey's	Telecommunications	StarHub, Optus
Breakfast cereal	Sanitarium	Aviation	Jetstar, JetBlue
Beverage	Nescafe, Bud Light	Casual Dining	Applebee's
Diapers	Huggies	Family Dining	IHOP
Ready to Eat Meals	Stouffers	Quick Serve Restaurants	KFC
Retailing		***Financial services***	
Gasoline	ExxonMobil, BP	Wealth management	Australian Super
Supermarkets	Coles	Insurance	Allianz
Discount Department	Kmart	***B2B***	
Durables		Pharmaceutical	Roche, GSK
Automobile	Honda	***Sundry***	
		Media	News Limited