



Norms for emotion-false memory lists

M. Chang¹ · C. J. Brainerd¹ · M. P. Toglia² · S. R. Schmidt³

Published online: 22 June 2020

© The Psychonomic Society, Inc. 2020

Abstract

False memory has been a flourishing research area for decades, and recently there has been considerable interest in how emotional content affects it. Literature reviews have noted a lack of normed materials that vary in emotional valence and arousal as a factor that contributes to the mixed findings on emotion-false memory effects. We report a pool of normed materials of this sort, the Cornell/Cortland Emotional Lists (CEL). This is a Deese/Roediger/McDermott (DRM) type list pool in which words' mean valence and arousal ratings are factorially manipulated across 32 lists. These lists' levels of mean backward associative strength (MBAS) are all high enough to induce significant levels of false memory. The lists were normed by administering them to 228 subjects at three different universities, all of whom responded to recall and recognition tests for the lists. The norming data revealed that false recall and false recognition were higher for negative lists than for positive lists, whereas true recall and true recognition were higher for positive lists than for negative lists. In addition, high arousal strengthened the valence effects on both true and false recall. These results indicate that the CEL lists are useful tools for emotion-false memory research.

Keywords Emotion · False memory · Valence · Arousal · DRM

False memory refers to memory for events that never actually happened. Although most instances of everyday false memory are innocuous, memory distortions can have serious consequences in some high-stakes situations in which memory accuracy is paramount, such as eyewitness identifications, police interrogations, and medical histories given during emergency treatments. An important objective of false memory research is to produce empirical findings and validate theoretical principles that can be applied to these high-stakes situations. For instance, police interrogation (e.g., Kassir et al. 2010), eyewitness interviews and identifications (e.g., Bruck and Ceci 1999; Poole and Lamb 1998; Wixted and Wells 2017), and

psychotherapy (e.g., Lindsay and Read 1994; Lynn et al. 2003) have been the foci of such work.

A hallmark of high-stakes remembering situations is that they generally involve emotion, either in the sense that the contexts are inherently emotional (e.g., police interviews) or in the sense that the events being remembered are emotional (e.g., violent crimes) or both (Bookbinder and Brainerd 2016; Brainerd and Reyna 2005). Naturally, that has stimulated research interest in how the emotional content of experience affects false memory.

Mixed findings in emotion-false memory research

In order to manipulate emotional content in experimentation, it is first necessary to agree upon the psychological dimensions of emotional content. Since Wundt (1912), it has been widely thought that at least two types of information are processed when we perceive emotional content in target materials—valence (positivity-negativity) and arousal (calming-exciting). In contemporary psychology, this idea was formalized in the circumplex model of emotion (Posner et al. 2005; Russell 1980), which posits that emotional experiences are mixtures of different values of valence and arousal.

Consequently, emotion-false memory research has revolved around the question of how false memory is influenced

This research was supported by National Institutes of Health Grant 1RC1AG036915 to the second author.

✉ M. Chang
mc2674@cornell.edu

¹ Department of Human Development and Human Neuroscience Institute, Cornell University, G331 MVR Hall, Ithaca, NY 14853, USA

² Department of Psychology, University of North Florida, Jacksonville, FL, USA

³ Department of Psychology, Middle Tennessee State University, Murfreesboro, TN, USA

by variations in the valence of target material and, to a far lesser extent, by variations in its arousal. Thus far, however, experimental findings about the valence effects on false memory remain elusive (for a review, see Bookbinder and Brainerd 2016). On the one hand, the results of some studies suggest that false memory levels are higher for emotionally valenced (positive/negative) materials than for comparable neutral materials (Bookbinder and Brainerd 2017; Brueckner and Moritz 2009; Gallo et al. 2009; Sharkawy et al. 2008). On the other hand, the results of other studies suggest that false memory levels are lower for emotionally valenced (positive/negative) materials than for neutral materials (Choi et al. 2013; Howe et al. 2010; Palmer and Dodson 2009; Pesta et al. 2001).

The exact causes of these inconsistent outcomes are not well understood, but a lack of methodological standardization across experiments is one likely contributor. A helpful first step in methodological standardization would be to develop a simple, uniform procedure that is easy to implement in a broad range of situations. An obvious candidate is a list-learning procedure that was originally devised by Deese (1959), which is currently the most widely used methodology in false memory research (Gallo 2006). In this procedure, subjects study lists of 12–15 related words (e.g., *bed, rest, awake, tired, dream, ...*). All of the list words are forward associates of a word that is not presented on the list (*sleep*), which is called the critical distractor or critical lure. Deese (1959) found that the critical distractors were falsely remembered at surprisingly high levels on immediate free recall tests. More than three decades later, Roediger and McDermott (1995) replicated the free recall results and found that the critical distractors were also falsely remembered at high levels on recognition tests.

Early DRM experiments in emotion-false memory research

In an early study, Budson et al. (2006) adapted the Deese/Roediger/McDermott (DRM; Deese 1959; Roediger and McDermott 1995) procedure by creating a set of emotional (negative) DRM lists, along with a parallel set of neutral DRM lists. The procedure of selecting words that are forward associates of a critical distractor was preserved, but the words on Budson et al.'s emotional lists had clear affective content (e.g., the *sick* list includes *cough, fever, ill, ...*), whereas those on their neutral lists did not (e.g., the *chair* list includes *table, sit, desk, ...*). Budson et al. found that false recognition of critical distractors was equivalent between the emotional and neutral lists. However, using a subset of Budson et al.'s lists, Howe (2007) and Howe et al. (2010) found that emotional lists produced higher false recognition and lower false recall than neutral lists. Sharkawy et al. (2008) replicated Howe et

al.'s finding for false recognition, but they found no difference in false recall between the two types of lists.

To understand these mixed findings, it is important to note that the Budson et al. (2006) methodology has two limitations. First, some lists in the neutral pool are not truly neutral, as their critical distractors receive highly positive ratings in norms for emotional words (e.g., *soft, girl* and *sweet*; Bookbinder and Brainerd 2016). According to Warriner et al.'s (2013; WKB) word norms, among Budson et al.'s 10 neutral critical distractors, 6 received mean valence ratings > 7 on a 9-point unhappy → happy scale. Thus, although Budson et al.'s (2006) emotional lists are clearly negative, their neutral lists are actually a mixture of neutral and positive words. The other limitation is that the emotional and neutral list pools confound valence with arousal. As Brainerd et al. (2008a) pointed out, although there is a valence difference such that the words on emotional lists are on average more negative than the words on neutral lists, the words on emotional lists are also on average more arousing than those on neutral lists. According to the WKB norms, the mean valence ratings (on a 9-point unhappy → happy scale) of Budson et al.'s (2006) emotional and neutral critical distractors are 2.69 and 6.47, respectively, and the corresponding mean arousal ratings (on a 9-point calm → exciting scale) are 5.59 and 3.48, respectively. Therefore, it is uncertain whether the memory effects that were produced by the two types of lists were due to the valence difference, the arousal difference, both, and/or the interaction between them.

Cornell/Cortland emotional lists

In order to overcome the limitations discussed above, Brainerd et al. (2008b) developed a new pool of emotional DRM lists, the Cornell/Cortland Emotional Lists (CEL; see Table 1). The CEL consists of 32 DRM lists, which are subdivided into four sets of lists: eight negative/high arousal lists, eight negative/low arousal lists, eight positive/high arousal lists and eight positive/low arousal lists. The CEL was initially constructed using the Nelson et al.'s (2004) norms of word association and the affective norms for English words (ANEW; Bradley and Lang 1999). The Nelson norms were used to isolate groups of words whose mean backward associative strength (MBAS; the mean probability for the list words to elicit the critical distractor as forward associates) is sufficiently high. This is critical for generating significant levels of false memory, because MBAS is the best-known predictor of the level of false memory that a DRM list induces (Hutchison and Balota 2005; Roediger et al. 2001). As shown in Table 2, the MBAS values of the four sets of lists are .26, .23, .21, .26, respectively, which are adequate to produce reliable levels of false recognition and false recall according to existing DRM norms (Roediger et al.

Table 1 The 32 Cornell/Cortland emotional lists and their critical distractors with mean valence, standard deviation valence, mean arousal, standard deviation arousal and backward association strength (BAS)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Neg/high											
Cold	4.32	2.11	3.55	2.57		Hurt	2.45	1.47	4.72	2.34	
Hot	5.73	1.83	4.76	2.66	.68	Pain	2.00	1.28	6.27	2.59	.55
Snow	6.78	1.78	4.57	2.72	.20	Cut	3.90	1.84	5.07	2.43	.06
Warm	7.50	1.92	3.35	2.50	.36	Ouch	--	--	--	--	.40
Winter	5.50	1.85	2.77	2.33	.28	Cry	3.22	2.41	5.45	2.82	.05
Ice	6.06	2.07	3.30	2.30	.36	Injury	2.32	1.63	5.56	2.33	.45
Wet	5.64	2.26	5.00	2.71	.11	Life	6.68	2.49	5.59	2.58	--
Frigid	3.77	1.93	3.87	2.36	.57	Sick	2.29	1.38	4.67	2.58	.02
Chilly	4.41	1.79	5.00	2.43	.40*	Fall	3.89	2.31	4.24	2.17	.05
Heat	4.77	2.62	5.75	2.69	.17	Harm	1.91	1.11	5.90	2.29	.64
Weather	6.05	1.59	4.21	2.32	.03	Hit	3.95	2.04	5.48	2.00	.07
Freeze	4.64	2.52	4.00	2.03	.46	Sore	2.80	1.20	4.06	2.26	.28
Air	6.71	1.85	3.25	2.55	--	Stop	4.73	1.16	4.14	2.50	.00
Shiver	3.61	2.12	4.65	2.54	.67	Wound	3.24	1.48	4.65	2.19	.22
Arctic	4.45	2.01	3.88	2.64	.64	Upset	2.45	1.05	4.49	2.67	.03
Flu	2.58	1.80	4.00	2.41	.37	Ache	3.27	2.01	4.30	2.26	.15
Mean	5.21	2.00	4.16	2.48	.38	Mean	3.33	1.67	4.99	2.41	.23
Thief	2.32	1.34	6.05	2.38		Spider	3.35	2.30	6.91	1.93	
Steal	2.17	1.15	5.30	2.79	.09	Web	5.68	1.67	4.14	2.83	.85
Robber	2.90	1.57	6.20	2.21	.36	Insect	4.43	1.93	4.67	2.59	.00
Crook	2.86	1.46	4.62	2.90	.46	Bug	3.45	1.47	6.06	2.39	.04
Burglar	2.67	1.91	5.32	2.78	.26	Fright	3.26	2.16	5.00	2.47	--
Money	7.10	2.27	6.86	2.49	.00	Fly	6.06	2.65	4.90	2.41	.00
Cop	4.50	2.33	4.90	2.65	--	Arachnid	--	--	--	--	.70*
Police	4.59	2.40	5.95	2.46	.00	Crawl	3.95	1.54	3.62	1.83	.00
Rob	2.10	1.21	5.54	2.86	.07	Tarantula	3.37	2.24	6.76	2.10	.74*
Jail	1.91	1.44	4.47	2.50	.01	Poison	2.16	1.46	6.01	2.31	--
Gun	3.66	2.55	7.74	1.91	--	Bite	3.52	2.09	5.10	2.31	--
Villain	3.00	1.82	4.91	2.76	--	Creepy	2.52	1.29	5.54	2.06	.06*
Crime	1.95	1.13	5.68	2.62	.03	Black widow	--	--	--	--	--
Bank	6.00	1.94	4.19	2.93	--	Monkey	5.82	2.13	5.15	2.50	.00
Bandit	3.88	1.98	5.18	2.77	.17	Feelers	--	--	--	--	.00*
Criminal	2.11	0.81	4.49	2.62	.05	Tiny	5.10	1.70	4.60	2.50	--
Mean	3.43	1.73	5.42	2.62	.17	Mean	4.11	1.86	5.13	2.36	.44
Fall	3.89	2.31	4.24	2.17		Anger	2.50	1.36	5.93	2.77	
Spring	7.64	1.46	5.50	2.57	.19	Mad	2.47	1.26	5.59	2.55	.39
Leaves	6.16	1.98	3.05	2.30	.17	Fear	2.93	1.79	6.14	2.76	.02
Winter	5.50	1.85	2.77	2.33	.00	Hate	1.96	1.33	6.26	2.31	.03
Down	--	--	--	--	.02	Rage	2.50	1.95	6.62	2.47	.54
Autumn	7.12	1.98	3.21	2.34	.53	Temper	3.89	1.76	4.90	2.07	.18
Hurt	2.45	1.47	4.72	2.34	.01	Fury	2.73	1.78	6.12	2.14	.31
Frosty	6.15	1.23	4.61	2.12	--	Ire	--	--	--	--	.18*
Summer	7.50	1.89	5.48	2.86	.02	Wrath	3.19	1.75	5.95	2.27	.13*
Catch	5.64	1.76	4.00	2.14	.03	Happiness	8.48	0.81	6.50	2.63	--
Cool	6.82	1.56	3.43	2.31	.00	Fight	3.54	1.99	6.33	2.40	.03

Table 1 (continued)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Trip	7.00	2.14	5.33	2.89	.29	Hatred	2.38	1.78	5.22	2.52	.07*
Drop	4.23	1.57	4.67	2.35	.23	Hostile	2.35	1.18	5.39	2.33	--
Rise	5.73	1.52	5.30	2.75	.33	Calm	6.89	2.00	1.67	1.91	--
Stand	5.27	1.83	3.10	2.05	.05	Emotion	6.62	1.88	4.75	2.79	--
Season	6.71	1.27	3.20	1.82	.19	Enrage	2.33	1.39	5.22	2.79	.24
Mean	5.99	1.68	4.17	2.37	.17	Mean	3.73	1.62	5.48	2.42	.22
Needle	3.97	1.73	4.36	2.35		Rough	3.68	1.69	5.43	1.97	
Thread	5.50	0.95	3.87	2.53	.76	Smooth	6.42	1.77	2.76	2.02	.42
Pin	5.29	1.52	3.73	1.98	.29	Bumpy	4.30	1.30	4.76	1.67	.15
Eye	6.18	1.47	3.95	2.27	.00	Road	5.59	1.53	3.81	2.16	--
Sewing	5.05	1.88	3.45	2.24	.18*	Tough	4.45	1.53	4.24	2.70	.19
Sharp	5.24	2.02	6.00	2.11	.03	Sandpaper	4.95	2.06	3.13	2.20	.43
Point	5.45	1.01	3.86	2.29	.02	Jagged	3.95	1.75	6.05	2.11	.13*
Prick	2.81	1.21	4.26	2.61	.11	Ready	6.64	1.97	3.90	2.43	.00
Thimble	--	--	--	--	.22	Coarse	4.50	1.58	3.66	2.23	.29*
Haystack	5.86	1.71	3.58	2.41	.42	Uneven	4.45	1.76	3.43	2.56	.02
Thorn	3.62	1.45	4.20	2.35	.03	Riders	5.36	1.22	4.41	2.46	.03*
Harm	1.91	1.11	5.90	2.29	--	Rugged	5.10	1.86	4.10	2.49	.17*
Injection	2.89	1.41	5.67	2.37	.33	Sand	6.16	1.92	3.43	2.6	--
Syringe	3.05	1.08	4.81	2.23	.52	Beard	5.09	2	3.18	2.08	.00
Cloth	6.20	1.44	2.58	1.77	--	Ground	5.28	1.56	2.35	1.73	--
Knitting	5.85	1.60	3.35	2.25	.14	Gravel	4.42	1.43	2.95	1.83	--
Mean	4.64	1.42	4.23	2.26	.26	Mean	5.11	1.68	3.74	2.22	.24
Neg/low											
Sick	2.29	1.38	4.67	2.58		Fat	2.74	1.74	3.89	2.50	
Ill	--	--	--	--	.82	Skinny	5.36	2.15	4.27	2.36	.42
Well	--	--	--	--	.17	Thin	5.47	1.58	4.50	2.98	.56
Cold	4.32	2.11	3.55	2.57	.05	Cat	6.95	1.85	4.50	2.48	.00
Hospital	7.27	1.72	3.33	2.24	.32	Ugly	2.47	1.93	4.43	2.18	.02
Vomit	1.98	1.48	4.82	2.46	.16	Lady	6.91	1.31	4.05	2.04	.00
Doctor	5.93	1.89	4.05	2.40	.05	Big	5.64	1.73	4.33	2.42	.00
Bed	7.16	1.86	3.00	2.32	.00	Cow	5.42	1.71	2.95	1.86	.01
Flu	2.58	1.80	4.00	2.41	.59	Slim	6.63	1.61	4.26	2.07	.42
Fever	2.92	1.93	5.02	2.54	.29	Albert	--	--	--	--	--
Nurse	5.41	2.25	4.64	2.82	--	Slob	2.48	1.33	4.70	2.43	.05
Medicine	5.90	2.39	4.00	1.99	.19	Obese	2.32	1.25	3.95	2.46	--
Nausea	1.68	1.00	4.59	2.52	.45	Man	5.42	1.22	4.36	2.49	.00
Virus	1.71	0.78	4.61	2.29	.35	Pig	4.83	2.20	3.68	1.73	.08
Pills	4.16	1.80	3.73	2.19	--	Diet	4.42	1.30	3.83	2.09	.19
Cure	7.81	1.33	4.50	2.76	.08	Large	5.77	1.90	5.35	2.41	.02
Mean	4.53	1.72	4.14	2.42	.29	Mean	5.01	1.65	4.23	2.29	.20
Shy	5.16	1.74	3.33	2.11		Dirt	4.50	1.79	3.44	2.13	
Quiet	6.47	1.58	1.95	1.58	.05	Mud	5.06	1.21	3.19	1.66	.60
Outgoing	6.89	1.20	5.71	2.10	.08	Sand	6.16	1.92	3.43	2.60	.03
Timid	3.37	1.95	3.15	1.84	.69	Clean	7.09	1.80	3.57	2.20	.02
Bashful	5.55	1.68	4.36	2.61	.73	Soil	5.21	1.55	2.76	1.76	.72
Scared	2.80	1.83	6.10	2.39	.00	Dirty	3.17	1.89	5.05	2.18	.00
Bold	6.68	1.49	4.86	2.42	.03	Filth	2.62	1.36	4.09	2.58	.69

Table 1 (continued)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Person	6.10	1.62	3.71	2.08	--	Ground	5.28	1.56	2.35	1.73	.18
Introvert	--	--	--	--	.18	Bath	7.00	1.75	3.25	2.35	.00
Modest	6.42	1.47	2.93	1.69	.31	Black	5.40	2.14	3.58	2.41	.00
Me	--	--	--	--	.00	Grime	--	--	--	--	.46
Loud	--	--	--	--	.00	Bike	6.10	2.28	3.62	2.37	.00
Afraid	2.25	1.16	5.12	2.57	.00	Brown	5.52	1.63	3.32	2.19	.01
Boring	2.71	1.68	2.85	1.87	.00	Grease	3.90	1.37	3.62	1.81	.00
Coy	5.50	1.32	3.72	2.05	--	Mess	3.50	1.29	4.70	2.43	.00
Extrovert	--	--	--	--	--	Unclean	2.95	1.72	4.12	2.56	--
Mean	4.98	1.54	4.04	2.11	.30	Mean	4.93	1.68	3.62	2.20	.34
Board	5.33	1.53	3.52	1.89		Trash	2.74	1.41	3.66	2.63	
Wood	5.82	1.47	3.50	1.70	.01	Garbage	2.88	1.42	3.84	2.36	.46
Chalk	5.00	1.38	2.90	1.92	.69	Waste	2.61	1.38	4.04	2.27	.07
Surf	6.05	1.67	5.57	1.96	.05	Can	6.41	1.53	3.14	1.93	.01
Room	5.55	1.53	3.10	1.58	.02	Refuse	3.42	1.35	4.32	2.06	.02
Game	6.79	1.23	5.05	2.27	.03	Sewage	2.33	1.56	4.09	2.37	.05*
Nail	4.60	1.70	3.05	2.46	.00	Bag	5.05	1.29	3.43	2.20	.00
Plane	5.72	1.85	4.91	2.18	.00	Junk	3.57	1.89	3.95	2.26	.13
School	5.41	2.26	4.57	2.36	.00	Rubbish	3.39	2.03	4.10	2.65	.40*
Ship	6.14	1.68	3.94	2.29	.00	Sweep	5.27	1.03	3.89	2.23	--
Walk	6.77	1.60	3.24	2.02	.00	Scraps	4.71	1.59	3.43	2.18	.05*
Two-by-four	--	--	--	--	--	Pile	4.74	1.37	4.43	2.13	.05
Black	5.40	2.14	3.58	2.41	.00	Dump	3.73	1.82	3.67	2.06	.22
Boat	6.36	1.40	4.05	2.18	.00	Landfill	3.14	1.71	3.65	2.06	.19*
Build	6.33	1.68	3.38	2.18	.00	Debris	4.20	1.74	4.23	2.29	.27*
Bulletin	5.60	1.67	2.95	2.01	.63	Litter	2.58	1.43	3.80	1.99	.21
Mean	5.82	1.66	3.84	2.11	.24	Mean	3.87	1.54	3.87	2.20	.14
Glass	5.48	1.67	3.14	1.88		Dead	2.02	1.39	3.91	2.15	
Window	6.47	1.62	3.27	1.91	.26	Alive	6.57	2.33	5.93	2.88	.55
Cup	5.94	2.04	2.60	2.01	.15	Gone	--	--	--	--	.01
Drink	6.67	1.83	5.19	2.56	.01	Cold	4.32	2.11	3.55	2.57	.00
Break	5.10	2.45	4.71	2.26	.10	Funeral	2.07	1.89	4.10	2.57	.14
Clear	6.14	1.70	2.71	1.71	.04	Asleep	6.50	1.89	2.00	2.43	--
Jar	5.71	1.82	2.77	1.88	.12	Bury	1.95	1.21	5.55	2.09	.22
Crystal	6.75	1.02	4.10	2.68	.15	End	4.73	1.75	3.95	2.40	.00
Mug	4.16	1.77	3.83	2.52	.14	Grave	2.40	2.09	4.54	2.02	.19
Bottle	5.47	0.84	3.32	1.84	.09	Sad	2.10	0.91	3.49	2.21	.00
Fragile	4.67	1.20	3.05	1.96	.10	Still	--	--	--	--	.00
Plastic	4.48	1.29	3.36	1.78	.03	Animal	7.06	1.95	4.30	2.70	.00
Sharp	5.24	2.02	6.00	2.11	.00	Cemetery	3.49	1.83	4.10	2.32	.21
Eye	6.18	1.47	3.95	2.27	.00	Deceased	2.73	2.07	4.86	2.94	--
House	7.19	1.40	3.95	2.65	.00	Die	1.67	1.65	6.90	2.21	.08
Juice	6.90	1.61	4.60	2.74	.00	Duck	6.11	1.20	4.00	1.84	.00
Mean	5.80	1.61	3.83	2.19	.11	Mean	3.98	1.76	4.41	2.40	.20
Pos/high											
God	5.90	2.19	5.56	2.83		Baby	6.67	2.36	4.97	2.65	
Jesus	--	--	--	--	.18	Child	7.20	1.47	5.33	2.59	.20
Heaven	7.50	2.24	4.23	2.86	.12	Cry	3.22	2.41	5.45	2.82	.11

Table 1 (continued)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Love	8.00	1.39	5.36	3.23	.00	Small	5.76	2.02	3.43	1.78	.00
Almighty	6.15	2.16	4.59	2.61	.62	Bottle	5.47	0.84	3.32	1.84	.02
Church	5.21	2.56	3.63	2.55	.32	Cute	7.56	1.07	5.05	2.82	.01
Lord	5.59	2.17	3.81	2.42	.47	Infant	6.65	1.93	4.10	2.72	.75
Religion	5.42	3.08	5.00	2.68	.19	Mother	7.53	1.86	4.73	2.92	.00
Creator	6.45	2.06	4.89	2.98	.50	Crib	6.43	2.34	4.26	2.73	.84
All-knowing	--	--	--	--	--	Doll	5.88	1.57	3.51	2.25	.13
Bible	5.74	2.77	3.65	2.25	.35	Little	5.91	1.66	4.20	1.96	.00
Damn	4.32	1.78	5.10	2.47	.06	Boy	5.84	1.70	4.11	2.49	.00
Devil	3.11	2.45	5.40	3.06	.09	Adult	5.90	2.10	4.36	2.50	.00
Faith	6.44	2.50	3.40	2.06	.20	Diaper	3.71	1.74	3.41	1.87	.59
Nun	5.79	2.53	2.43	1.88	--	Girl	7.15	1.56	5.23	2.68	.00
Holy	6.06	2.65	3.37	2.87	.09	Young	6.31	1.59	4.09	2.22	.00
Mean	5.83	2.33	4.22	2.61	.27	Mean	6.03	1.72	4.31	2.41	.33
Love	8.00	1.39	5.36	3.23		Music	7.67	1.82	5.57	3.09	
Hate	1.96	1.33	6.26	2.31	.41	Note	5.78	1.73	3.33	2.06	.13
Kiss	7.78	1.64	6.05	3.03	.23	Sound	6.45	1.34	4.14	2.35	.21
Like	7.44	1.46	4.40	2.50	.34	Piano	6.40	1.76	3.61	2.10	.23
Happy	8.47	1.28	6.05	2.13	.00	Sing	7.50	1.37	4.10	2.49	.03
Heart	6.95	1.87	5.07	2.96	.15	Radio	6.00	2.21	3.84	2.57	.27
Care	7.64	1.29	2.67	1.85	.26	Band	6.44	1.62	4.52	2.93	.43
Admire	7.35	1.27	5.00	2.27	.09	Melody	6.65	1.64	4.86	2.67	.24
Adore	7.37	1.38	5.96	2.33	.57	Horn	5.00	1.38	4.22	2.58	.01
Close	5.23	2.09	3.81	2.06	.00	Concert	7.00	1.22	5.17	3.14	.40
Friendship	7.76	1.64	5.20	2.98	.16	Instrument	6.68	1.67	3.88	2.60	.15
Girl	7.15	1.56	5.23	2.68	.00	Symphony	7.15	2.18	4.19	2.56	.33
Happiness	8.48	0.81	6.50	2.63	.10	Jazz	6.57	1.86	6.05	2.76	.37
Hug	8.00	1.39	5.36	3.23	.13	Orchestra	6.50	1.87	4.07	2.63	.31
Kindness	1.96	1.33	6.26	2.31	.08	Art	6.85	1.89	4.48	2.48	.02
Life	7.78	1.64	6.05	3.03	.02	Rhythm	6.42	1.17	5.14	2.54	.28
Mean	6.75	1.47	5.32	2.55	.21	Mean	6.49	1.66	4.37	2.56	.23
City	6.12	1.54	5.08	2.54		Beach	7.21	1.51	5.10	2.23	
Town	5.59	1.50	3.81	2.06	.53	Sand	6.16	1.92	3.43	2.60	.72
Crowded	3.20	1.24	3.95	2.50	.00*	Sun	6.92	1.89	4.64	2.54	.06
State	5.73	1.32	3.29	1.85	.12	Ocean	7.39	2.01	3.50	2.83	.08
Capital	5.29	1.82	3.14	2.08	.10	Water	7.00	1.74	3.71	2.39	.00
Streets	5.07	1.00	3.45	2.18	.00	Towel	6.14	1.53	2.90	2.07	.01
Subway	5.44	2.12	4.41	2.70	--	Ball	6.14	1.39	3.48	2.02	.00
Country	6.14	1.67	3.71	2.47	.07	Summer	7.50	1.89	5.48	2.86	.07
New York	--	--	--	--	.38	Girls	7.15	1.56	5.23	2.68	.00
Village	5.95	1.79	3.40	2.26	.02	Swim	6.71	2.03	6.05	3.05	.00
Metropolis	4.95	1.96	4.95	2.48	.54*	Fun	8.37	0.96	6.32	2.62	.05
Big	5.64	1.73	4.33	2.42	.00	Wave	6.32	1.00	4.19	2.16	.09
Chicago	--	--	--	--	.15*	Blanket	7.05	1.54	2.23	1.74	.02
Suburb	5.52	1.57	4.25	2.24	.27	Bum	3.64	2.24	4.39	2.50	.05
County	5.18	1.79	3.40	2.16	.20	Coast	5.90	2.28	4.17	2.64	.10
Urban	4.95	2.66	4.41	2.84	.36*	Shore	--	--	--	--	.46
Mean	5.28	1.71	3.88	2.33	5.28	Mean	6.60	1.71	4.27	2.48	.16

Table 1 (continued)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Hug	8.23	0.87	4.85	2.95		Pretty	7.70	1.45	5.50	2.33	
Kiss	7.78	1.64	6.05	3.03	.22	Ugly	2.47	1.93	4.43	2.18	.29
Love	8.00	1.39	5.36	3.23	.01	Girl	7.15	1.56	5.23	2.68	.03
Squeeze	4.85	1.46	3.91	2.52	.26	Beautiful	7.61	1.74	5.71	2.82	.39
Embrace	7.00	1.45	5.82	2.58	.59	Cute	7.56	1.07	5.05	2.82	.20
Affection	7.89	1.05	5.64	2.63	.04	Adorable	7.40	1.60	4.66	2.79	.04
Hold	6.05	1.76	4.86	2.17	.07	Attractive	7.19	1.91	4.91	2.94	.35
Cuddle	7.60	0.99	3.83	2.96	.20	Flower	7.30	1.87	3.67	2.59	.07
Arms	5.44	1.70	3.44	2.05	.02	Gorgeous	7.57	1.66	4.92	2.41	.06
Bear	5.33	2.25	4.50	3.07	.03	Hot	5.73	1.83	4.76	2.66	.00
Caress	6.82	1.45	4.60	2.70	.03	Model	6.38	1.99	3.90	2.61	.12
Tight	--	--	--	--	.00	Pink	6.68	1.89	4.35	2.98	.09
Fondle	4.32	2.36	5.18	2.50	--	Appealing	--	--	--	--	--
Warmth	7.53	1.38	4.33	2.35	.00	Sweet	7.77	1.38	4.14	2.92	--
Clinch	4.85	1.93	4.00	2.30	--	Charming	7.05	2.15	5.00	2.79	--
Clutch	4.80	1.70	4.04	1.90	--	Lovely	7.55	1.65	3.95	2.75	--
Mean	6.30	1.61	4.68	2.57	.15	Mean	6.82	1.73	4.62	2.71	.16
Pos/low											
Soft	7.13	1.66	3.04	2.46		Green	6.29	2.33	4.07	2.59	
Hard	4.35	1.97	4.50	2.75	.56	Grass	6.47	1.26	3.39	2.35	.36
Light	6.55	1.82	3.52	2.58	.00	Yellow	6.09	1.82	3.83	2.33	.18
Pillow	7.00	1.87	2.90	2.21	.24	Blue	6.53	2.34	3.69	2.40	.14
Plush	6.32	2.16	3.05	2.17	.18	Color	7.05	2.22	3.18	1.62	.04
Loud	--	--	--	--	.33	Red	5.67	2.07	5.02	2.56	.02
Cotton	6.05	1.96	2.48	2.11	.17	Leaf	6.16	1.98	3.05	2.30	.04
Fur	5.65	1.90	3.64	2.26	.06	Arrow	5.24	1.81	3.91	2.59	.00
Touch	6.64	1.76	4.90	2.14	.06	Clover	5.37	1.50	3.06	2.04	--
Fluffy	7.11	1.29	3.45	2.11	.27	Frog	5.84	1.34	4.07	2.61	.08
Warm	7.50	1.92	3.35	2.50	.06	Vegetables	6.79	1.55	3.75	2.83	.10
Furry	6.63	2.24	3.70	2.08	.06	Apple	6.62	1.72	3.52	2.06	.03
Downy	5.77	1.97	3.11	1.66	--	Bean	6.00	2.08	2.95	2.46	.15
Kitten	7.58	1.72	3.19	2.66	.03	Elf	6.10	2.17	4.30	2.14	.05
Skin	5.78	1.63	3.25	2.51	.16	Go	6.32	1.25	4.86	2.41	.03
Tender	6.47	1.75	3.22	2.21	.30	Dollar	7.39	1.51	5.57	2.73	.03
Mean	6.39	1.85	3.45	2.28	.19	Mean	6.24	1.77	3.88	2.36	.10
Nice	6.95	2.04	3.53	2.44		Sleep	7.22	1.65	3.60	2.57	
Mean	2.43	1.69	4.81	2.44	.21	Bed	7.16	1.86	3.00	2.32	.64
Sweet	7.77	1.38	4.14	2.92	.10	Rest	7.86	1.25	2.29	2.35	.48
Good	7.89	1.24	3.66	2.72	.01	Awake	6.76	1.48	5.50	2.67	.00
Bad	3.24	1.92	4.86	2.69	.00	Tired	4.29	1.35	3.67	2.18	.48
Kind	7.78	1.12	3.19	2.02	.37	Dream	7.43	1.80	4.37	2.81	.25
Pleasant	7.24	1.51	2.91	2.52	.16	Wake	6.57	1.96	3.80	2.61	.28
People	5.70	2.01	3.77	2.34	.02	Snooze	6.36	1.47	4.00	2.49	.52
Friendly	7.84	0.96	4.27	2.81	.29	Pillow	7.00	1.87	2.90	2.21	.32
Smile	7.89	2.19	4.62	3.09	.00	Doze	5.95	1.43	3.37	2.59	.68
Happy	8.47	1.28	6.05	2.13	--	Slumber	6.24	1.70	2.73	2.14	.51
Friend	6.79	2.49	4.29	2.69	.01	Snore	3.00	1.25	3.71	1.98	.44
Pretty	7.70	1.45	5.50	2.33	.03	Nap	6.53	1.95	4.14	2.41	.73

Table 1 (continued)

CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS	CDs/Ts	Mean valence	SD valence	Mean arousal	SD arousal	BAS
Nasty	2.95	2.14	5.44	2.23	.05	Peace	7.75	1.50	4.65	2.77	--
Delightful	7.38	1.83	6.16	2.75	--	Yawn	5.62	1.63	2.84	1.93	.24
Polite	6.57	1.93	2.95	1.86	.36	Drowsy	4.25	2.10	2.83	1.87	.55
Mean	6.51	1.68	4.44	2.50	.15	Mean	6.18	1.64	3.59	2.36	.47
Bird	6.75	1.78	3.83	2.48		Grass	6.47	1.26	3.39	2.35	
Nest	5.65	1.27	3.35	2.25	.68	Green	6.29	2.33	4.07	2.59	.25
Parrot	6.79	1.36	4.65	2.41	.63	Cut	3.90	1.84	5.07	2.43	.00
Eagle	6.47	1.50	4.57	2.77	.49	Lawn	6.05	1.84	2.62	1.63	.40
Feather	6.30	1.53	3.29	2.10	.36	Mow	4.94	1.86	4.17	2.39	.28
Animal	7.06	1.95	4.30	2.70	.02	Hopper	--	--	--	--	--
Wings	6.15	1.69	4.32	2.63	.48	Pot	5.81	2.09	4.00	2.00	.00
Canary	6.37	1.95	3.50	1.85	.69	Mower	4.76	1.04	2.79	2.59	.22
Cage	3.11	1.41	4.20	2.53	.17	Weed	4.65	2.10	4.78	2.81	.45
Bath	7.00	1.75	3.25	2.35	.00	Grow	6.28	2.08	2.60	1.73	.00
Dog	7.00	2.07	5.43	2.64	.00	High	5.77	1.77	4.10	2.63	.00
Flight	6.11	1.97	4.20	2.48	.02	Roots	5.33	1.88	3.62	2.31	--
Song	7.59	1.76	4.60	2.58	.03	Smoke	3.44	2.23	5.00	2.70	.00
Bluejay	6.64	1.50	3.43	2.27	.74	Turf	4.62	2.09	3.95	2.24	--
Tree	7.59	1.69	2.67	2.29	.00	Yard	5.70	1.92	3.68	2.03	.24
Robin	6.63	2.17	2.64	2.24	.63	Meadow	7.30	1.53	2.62	2.16	.22
Mean	6.43	1.70	3.89	2.41	.41	Mean	5.35	1.90	3.79	2.30	.29
Fish	6.42	1.67	3.33	2.51		River	6.72	1.72	4.22	2.86	
Water	7.00	1.74	3.71	2.39	.00	Water	7.00	1.74	3.71	2.39	.00
Swim	6.71	2.03	6.05	3.05	.01	Stream	6.90	2.00	4.35	2.78	.32
Sea	6.56	2.09	2.80	2.33	.03	Lake	7.13	1.34	2.64	2.08	.14
Scales	5.06	1.51	3.80	2.48	.29	Mississippi	--	--	--	--	.65*
Smell	6.39	1.79	5.24	2.83	.01	Boat	6.36	1.40	4.05	2.18	.00
Cod	5.21	1.81	3.95	2.44	.58	Tide	6.55	1.70	5.47	2.89	--
Food	7.52	1.38	4.69	2.63	.00	Canoe	5.76	2.02	3.70	2.34	.08
Trout	5.62	1.88	3.85	2.39	.91	Flow	5.68	0.89	3.71	2.28	.28
Gold	7.28	1.83	6.35	2.64	.00	Run	6.41	1.40	5.24	2.68	.00
Ocean	7.39	2.01	3.50	2.83	.00	Barge	5.44	1.69	3.15	2.21	.05*
Fry	5.95	2.31	5.00	2.97	.09	Creek	6.32	1.60	2.67	1.57	.40
Shark	4.02	1.80	5.27	2.65	.14	Brook	7.00	1.38	3.33	2.54	.16
Tank	4.71	1.58	4.32	2.68	.12	Bend	5.75	0.91	4.40	2.28	.00
Bass	5.23	1.38	4.87	2.69	.32	Bridge	5.44	1.42	3.48	2.73	.20
Bowl	5.67	1.49	3.21	2.20	.01	Wind	5.67	2.22	3.70	2.47	.00
Mean	6.02	1.78	4.44	2.61	.23	Mean	6.24	1.55	3.83	2.39	.23

Note. *SD* standard deviation, *CD* critical distractors, *T* targets, *Neg* negative, *Pos* positive. The critical distractor of each list is shown in bold font. The mean valence, SD valence, mean arousal and SD arousal are drawn from the WKB norm (Warriner et al., 2013), and the BAS data was drawn from the norming data published in Nelson et al., (2004). For some words whose BAS values are missing in the Nelson norm, we used the BAS values reported in Roediger et al. (2001) instead. Such values are obtained in a norming procedure similar to Nelson et al.'s method, and they are indicated with an asterisk in the current table. We use "--" to indicate the scores for words that are missing from the data sources mentioned above. The mean values for the mean valence, SD valence, mean arousal, SD arousal and BAS of the list words are also displayed for each list

2001). The ANEW norms were used to identify words that vary in valence and arousal. When the lists were first

constructed, many of the words had not been rated in the ANEW norms, and hence, their valence and arousal ratings

Table 2 Valence-arousal counterbalancing and BAS of the Cornell/Cortland emotional lists

Valence/arousal	Mean valence		SD valence		Mean arousal		SD arousal		MBAS
	Ts	CD	Ts	CD	Ts	CD	Ts	CD	
Neg/high	4.44	3.31	1.71	1.79	4.66	5.15	2.39	2.31	0.26
Neg/low	4.79	3.78	1.64	1.58	3.98	3.70	2.24	2.23	0.23
Pos/high	6.32	7.19	1.74	1.64	4.51	5.25	2.54	2.73	0.21
Pos/low	6.21	6.74	1.74	1.76	3.89	3.63	2.41	2.53	0.26

Note. *Ts* the 15 targets on each DRM list; *CD* the critical distractor of each DRM list. *BAS* backward associative strength from the 15 targets to the critical distractor of each list

were unknown. Eventually, however, Warriner et al. (2013) published a follow-up of the ANEW norms that included 13 times as many words. We have taken advantage of the Warriner et al. (WKB) norms to identify the mean valence and arousal scores for nearly all of the CEL list words and all of the critical distractors in Table 1.

The most important feature of the CEL is that the valence and arousal scores for both critical distractors and list words are counterbalanced across the four sets of lists (see Table 2), allowing researchers to disentangle the effects of valence and arousal on both true memory (recall/recognition of list words) and false memory (recall/recognition of critical distractors). The 2 (valence: positive, negative) \times 2 (arousal: high, low) analyses of variance (ANOVAs) revealed that the mean valence of the list words is lower for the 16 negative lists than for the 16 positive lists, $F(1, 28) = 42.72$, $MSE = .46$, $p < .0001$, partial $\eta^2 = .60$. The mean arousal of the list words is higher for the 16 high arousal lists than for the 16 low arousal lists, $F(1, 28) = 18.24$, $MSE = .17$, $p = .0002$, partial $\eta^2 = .39$. Also, the mean valence of the critical distractors is lower for the 16 negative lists than for the 16 positive lists, $F(1, 28) = 102.25$, $MSE = .91$, $p < .0001$, partial $\eta^2 = .79$. The mean arousal of the critical distractors is higher for the 16 high arousal lists than for the 16 low arousal lists, $F(1, 28) = 37.36$, $MSE = .50$, $p < .0001$, partial $\eta^2 = .57$.

Emotional ambiguity hypothesis

Recent research on the cognitive effects of the valence and arousal has stressed the importance of variability in items' average valence and arousal intensity ratings. Specifically, according to the emotional-ambiguity hypothesis (Brainerd 2018; Mattek et al. 2017), the relation between valence and arousal intensity and their effects on other psychological processes depend upon a variable that has been neglected in prior research: valence ambiguity. Valence ambiguity is the level of uncertainty in people's subjective perception of an item's valence, which can be measured with the standard deviation (SD) of items' valence ratings in normed materials.

Extensive data, using different stimuli (e.g., words, pictures), have demonstrated that (a) items' average valence intensity ratings are unrelated to the uncertainty of those ratings (Brainerd 2018), and (b) valence and arousal intensity ratings are strongly related (arousal ratings increase as valence ratings increase) only when the uncertainty of valence ratings is low (Brainerd 2018; Brainerd and Bookbinder 2019; Mattek et al. 2017). Thus, it is possible that the ambiguity of emotional content, as well as its intensity, determines how memory is affected by emotional content.

The fact that the emotional ambiguity hypothesis and related findings are so recent means that ambiguity of valence and arousal ratings has not figured in prior emotion-false memory experiments. In order to encourage such research, apart from mean valence and arousal intensity ratings, the descriptive data of the CEL in Table 1 also includes SDs of those ratings.

The present study

The purpose of the present study was to generate norming data for the CEL that investigators can use to select lists for emotion-false memory experiments. Using procedures that are very similar to those used in norming studies for other pools of DRM lists (e.g., Roediger et al. 2001; Stadler et al. 1999), we collected data from three different universities and tested both recall and recognition of all 32 CEL lists. All subjects were presented with 16 of the 32 lists, followed by free recall and then recognition tests. Hence, variability in true recall/recognition and false recall/recognition could all be tracked as a function of the valence and arousal of list words and critical distractors.

Method

Subjects

The sample consisted of 228 undergraduates in total—109 undergraduates (73 females and 36 males) were recruited from University A, which is a private university in the northeastern

US, and 57 undergraduates (30 females and 27 males) were recruited from University B, which is a public university in the northeastern US. Sixty-two undergraduates (32 females and 30 males) were recruited from University C, which is a public university in the southeastern US. The subjects at all three universities participated in order to fulfill a course requirement. All subjects were native English speakers. The data were gathered over a 3-year period, with the subjects in the different universities participating during different academic years.

Materials

The materials were the 32 CEL lists shown in Table 1. The critical distractor and list words for each list as well as the mean valence, SD valence, mean arousal, SD arousal, and BAS data for each word are displayed in Table 1. As mentioned, the valence and arousal ratings were drawn from the WKB norms (Warriner et al. 2013), and the BAS data were drawn from the Nelson et al.'s (2004) norms of word association. The 32 lists were evenly divided into negative/high arousal lists, negative/low arousal lists, positive/high arousal lists and positive/low arousal lists (see Table 2).

Procedure

Overall, the procedure resembled the DRM norming methodology of Stadler et al. (1999) and Roediger et al. (2001). Subjects were tested in small groups of 4–5 individuals. Before the experiment began, all subjects were provided with a recall booklet. For each subject, sixteen 15-word DRM lists were presented via computer recordings. The words on individual lists were presented at a 2-s rate in a neutral voice, with a 2-s interval between consecutive words. The 16 lists that were administered to individual subjects were counterbalanced across subjects, so that each list was presented to an equal number of subjects. After each list was presented, subjects performed written free recall for that list before the next list was presented. Subjects were given 90 s to recall each list, and they were told to stop writing and turn to the next page of the recall booklet when time was up. Following the presentation of all 16 lists, subjects worked on math problems for 1 min as a buffer task. After that, they completed a self-paced old/new recognition test with a scantron. There were 74 test items on the recognition test, including 48 targets (three list words from each list, taken from the 5th, 7th, and 9th list presentation positions), 16 critical distractors (one critical distractor from each list), and ten unrelated distractors (randomly selected from unrepresented lists). For each test item, subjects were instructed to bubble “a” if it was an old word and bubble “b” if it was a new word.

Results

The descriptive data for recall and recognition are summarized in Table 3 and Table 4, respectively. In Table 3, true and false recall results from the three universities (A, B, and C) and the grand means are reported. The true recall level was calculated as the number of targets recalled divided by the total number of targets, and the false recall level was calculated as the number of critical distractors recalled divided by the total number of critical distractors. The recall data for the *spider* list and the *fall* list are missing from the University A protocols because these two lists were inadvertently not presented to those subjects.

In Table 4, true recognition was measured by the bias-corrected acceptance rate for targets, and false recognition was measured by the bias-corrected acceptance rate for critical distractors. Here, we used the two-high-threshold statistic *Pr* (Snodgrass and Corwin 1988) for bias correction, which is calculated by subtracting false alarm rates for unrelated distractors from hit rates for targets and from false alarm rates for critical distractors. The University A false recognition data are missing for the *Pretty list* because its critical distractor had been presented as part of another list during the study phase. For the same reason, the University C false recognition data are missing for the *Hurt list*, *Spider list*, *Sick list* and *Fall list*. Importantly, ample false recognition data for these items are available from the other two universities.

We conducted 2 (valence: positive, negative) \times 2 (arousal: high, low) repeated measures analyses of variance (ANOVAs), using false recall, false recognition, true recall, and true recognition as dependent variables, respectively. In addition, we conducted post hoc tests using Tukey's honest significant difference test (HSD). The ANOVAs and post hoc analyses were conducted for the aggregated data of the three universities and separately for the University A, B, and C data, using the *lmerTest* package in R (Kuznetsova et al. 2017). In what follows, we first report the false recall and false recognition results as emotion-false memory is the topic of primary interest. Next, we report the true recall and true recognition results. All results that are reported below were significant at or beyond the .05 level of confidence.

False recall

Aggregated results

The aggregated results for the three universities produced a valence main effect, $F(1, 684) = 51.27$, $MSE = .04$, $p < .0001$, partial $\eta^2 = .07$, and a Valence \times Arousal interaction, $F(1, 684) = 36.38$, $MSE = .04$, $p < .0001$, partial $\eta^2 = .05$. According to the post hoc tests, the false recall levels were higher for negative lists than for positive lists. The interaction indicated that

Table 3 Mean true and false recall

Valence/arousal	University A		University B		University C		Grand means	
	True recall	False recall	True recall	False recall	True recall	False recall	True recall	False recall
Negative/high								
Cold	.57 (.13)	.33 (.48)	.63 (.11)	.48 (.51)	.56 (.18)	.42 (.05)	.59	.41
Hurt	.57 (.13)	.26 (.43)	.50 (.13)	.27 (.45)	.59 (.15)	.29 (.46)	.55	.27
Thief	.60 (.11)	.23 (.42)	.62 (.13)	.33 (.48)	.62 (.11)	.39 (.50)	.61	.32
Spider	--	--	.62 (.12)	.21 (.42)	.63 (.18)	.16 (.37)	.63	.19
Fall	--	--	.65 (.13)	.30 (.47)	.61 (.23)	.26 (.44)	.63	.28
Anger	.48 (.13)	.31 (.47)	.48 (.13)	.44 (.51)	.50 (.23)	.26 (.44)	.49	.34
Needle	.57 (.13)	.48 (.51)	.59 (.15)	.66 (.48)	.59 (.19)	.65 (.49)	.58	.60
Rough	.56 (.15)	.45 (.51)	.58 (.13)	.86 (.35)	.57 (.16)	.68 (.48)	.57	.66
Grand means	.56	.34	.58	.44	.58	.39	.58	.39
Negative/low								
Sick	.60 (.12)	.08 (.28)	.60 (.18)	.26 (.45)	.65 (.15)	.03 (.18)	.62	.12
Fat	.66 (.12)	.32 (.48)	.63 (.15)	.31 (.47)	.61 (.17)	.42 (.50)	.63	.35
Shy	.59 (.14)	.40 (.49)	.50 (.13)	.34 (.48)	.52 (.14)	.61 (.50)	.54	.45
Dirt	.59 (.15)	.28 (.45)	.56 (.15)	.41 (.50)	.59 (.20)	.16 (.37)	.58	.28
Board	.51 (.13)	.15 (.33)	.52 (.11)	.24 (.44)	.52 (.17)	.19 (.40)	.52	.19
Trash	.56 (.15)	.28 (.46)	.53 (.14)	.21 (.41)	.56 (.10)	.55 (.51)	.55	.35
Glass	.58 (.11)	.38 (.49)	.59 (.15)	.37 (.49)	.61 (.21)	.32 (.48)	.59	.36
Dead	.57 (.13)	.16 (.37)	.55 (.14)	.11 (.32)	.60 (.20)	.10 (.30)	.57	.12
Grand means	.58	.26	.56	.28	.58	.30	.58	.28
Positive/high								
God	.63 (.14)	.13 (.33)	.66 (.14)	.07 (.27)	.72 (.14)	.13 (.34)	.67	.11
Baby	.67 (.16)	.03 (.17)	.66 (.13)	.31 (.47)	.66 (.21)	.26 (.44)	.66	.20
Love	.60 (.17)	.45 (.51)	.53 (.10)	.34 (.48)	.56 (.13)	.32 (.48)	.56	.37
Music	.62 (.13)	.16 (.37)	.59 (.13)	.28 (.45)	.58 (.20)	.13 (.34)	.60	.19
City	.67 (.14)	.29 (.46)	.60 (.08)	.48 (.51)	.68 (.20)	.39 (.50)	.65	.39
Beach	.71 (.13)	.12 (.32)	.66 (.15)	.31 (.47)	.75 (.12)	.16 (.37)	.71	.20
Hug	.54 (.16)	.14 (.35)	.56 (.14)	.19 (.40)	.60 (.22)	.06 (.25)	.57	.13
Pretty	.73 (.11)	.23 (.42)	.71 (.11)	.11 (.32)	.67 (.21)	.13 (.34)	.70	.16
Grand means	.65	.19	.62	.26	.65	.20	.64	.22
Positive/low								
Soft	.56 (.12)	.28 (.45)	.52 (.13)	.41 (.50)	.54 (.20)	.35 (.49)	.54	.35
Green	.61 (.15)	.40 (.50)	.63 (.10)	.37 (.49)	.66 (.18)	.29 (.46)	.63	.35
Nice	.57 (.13)	.17 (.34)	.57 (.13)	.21 (.41)	.57 (.19)	.35 (.49)	.57	.24
Sleep	.64 (.14)	.30 (.47)	.53 (.15)	.59 (.50)	.56 (.19)	.45 (.51)	.58	.45
Bird	.62 (.17)	.08 (.28)	.62 (.12)	.22 (.42)	.61 (.30)	.06 (.25)	.62	.12
Grass	.56 (.11)	.25 (.44)	.62 (.15)	.28 (.45)	.62 (.12)	.19 (.40)	.60	.24
Fish	.62 (.15)	.13 (.33)	.59 (.14)	.22 (.42)	.66 (.22)	.16 (.37)	.62	.17
River	.61 (.16)	.29 (.46)	.61 (.16)	.28 (.45)	.58 (.20)	.26 (.44)	.60	.28
Grand means	.60	.24	.59	.32	.60	.26	.60	.27

Note. The standard deviations are included in the parentheses

negative/high arousal lists elicited more false recall than positive/high arousal lists, $p < .0001$, but false recall levels for negative/low arousal lists and that for positive/low arousal lists did not differ.

University A

There was a main effect of valence, $F(1, 331) = 28.49$, $MSE = .04$, $p < .0001$, partial $\eta^2 = .08$, and a Valence \times Arousal

Table 4 Mean bias-corrected true and false recognition

Valence/ arousal	University A		University B		University C		Grand means	
	True recognition	False recognition	True Recognition	False recognition	True recognition	False recognition	True recognition	False recognition
Negative/high								
Cold	.60 (.31)	.65 (.44)	.58 (.23)	.37 (.46)	.52 (.26)	.37 (.50)	.57	.46
Hurt	.47 (.36)	.60 (.49)	.09 (.37)	.44 (.52)	.58 (.33)	--	.38	.52
Thief	.70 (.30)	.59 (.48)	.50 (.49)	.54 (.57)	.59 (.28)	.49 (.48)	.60	.54
Spider	.57 (.34)	.12 (.40)	.05 (.21)	.00 (.12)	.41 (.49)	--	.34	.06
Fall	.71 (.27)	.33 (.49)	.54 (.40)	.47 (.56)	.44 (.28)	--	.56	.40
Anger	.58 (.33)	.77 (.38)	.32 (.31)	.65 (.46)	.40 (.34)	.62 (.40)	.43	.68
Needle	.65 (.31)	.56 (.51)	.54 (.44)	.54 (.35)	.50 (.21)	.60 (.37)	.56	.57
Rough	.49 (.31)	.79 (.39)	.20 (.24)	.68 (.37)	.28 (.25)	.70 (.25)	.32	.72
Grand means	.60	.55	.35	.46	.47	.56	.47	.49
Negative/low								
Sick	.52 (.34)	.41 (.50)	.46 (.36)	.54 (.46)	.34 (.30)	--	.44	.48
Fat	.68 (.30)	.44 (.48)	.60 (.30)	.30 (.49)	.72 (.26)	.37 (.51)	.67	.37
Shy	.65 (.30)	.65 (.47)	.30 (.35)	.33 (.57)	.53 (.26)	.66 (.43)	.49	.55
Dirt	.70 (.28)	.74 (.39)	.37 (.33)	.65 (.47)	.40 (.31)	.39 (.48)	.49	.59
Board	.44 (.33)	.36 (.52)	.21 (.37)	.23 (.48)	.45 (.34)	.47 (.50)	.37	.35
Trash	.57 (.34)	.50 (.49)	.51 (.29)	.61 (.37)	.63 (.29)	.60 (.46)	.57	.57
Glass	.60 (.47)	.58 (.49)	.40 (.42)	.42 (.35)	.29 (.33)	.45 (.44)	.43	.48
Dead	.60 (.26)	.57 (.49)	.42 (.35)	.47 (.53)	.29 (.28)	.46 (.49)	.44	.50
Grand means	.60	.53	.41	.44	.46	.49	.49	.49
Positive/high								
God	.76 (.23)	.28 (.49)	.55 (.46)	.25 (.56)	.45 (.22)	.10 (.51)	.59	.21
Baby	.55 (.35)	.30 (.49)	.43 (.25)	.13 (.48)	.48 (.27)	.25 (.51)	.49	.23
Love	.55 (.33)	.33 (.48)	.39 (.30)	.37 (.48)	.25 (.51)	.24 (.51)	.40	.31
Music	.64 (.30)	.25 (.48)	.40 (.34)	.30 (.55)	.43 (.29)	.00 (.44)	.49	.18
City	.69 (.26)	.59 (.48)	.53 (.39)	.54 (.49)	.31 (.32)	.16 (.50)	.51	.43
Beach	.71 (.27)	.45 (.49)	.51 (.28)	.44 (.47)	.41 (.28)	.18 (.51)	.54	.36
Hug	.59 (.31)	.50 (.50)	.55 (.36)	.50 (.49)	.33 (.29)	.32 (.44)	.49	.44
Pretty	.79 (.20)	--	.63 (.43)	.54 (.50)	.46 (.22)	.13 (.51)	.63	.34
Grand means	.66	.39	.50	.38	.39	.17	.52	.31
Positive/low								
Soft	.59 (.35)	.63 (.47)	.46 (.32)	.68 (.43)	.64 (.26)	.58 (.46)	.59	.66
Green	.55 (.31)	.46 (.51)	.51 (.37)	.25 (.55)	.69 (.30)	.26 (.50)	.61	.35
Nice	.62 (.31)	.65 (.47)	.42 (.30)	.37 (.48)	.79 (.19)	.81 (.37)	.62	.62
Sleep	.77 (.30)	.30 (.51)	.46 (.34)	.68 (.33)	.75 (.27)	.55 (.50)	.68	.52
Bird	.69 (.31)	.15 (.45)	.53 (.41)	.29 (.54)	.62 (.26)	.19 (.48)	.64	.24
Grass	.63 (.31)	.32 (.48)	.59 (.29)	.37 (.46)	.79 (.26)	.39 (.50)	.68	.37
Fish	.49 (.30)	.28 (.46)	.29 (.39)	.32 (.48)	.47 (.33)	.29 (.50)	.45	.33
River	.60 (.40)	.50 (.48)	.47 (.28)	.27 (.48)	.66 (.28)	.36 (.50)	.59	.39
Grand means	.62	.41	.47	.40	.68	.43	.59	.41

Note. The standard deviations are included in the parentheses

interaction, $F(1, 331) = 12.62$, $MSE = .04$, $p = .0004$, partial $\eta^2 = .04$. Post hoc tests suggested that false recall was higher for

negative lists than for positive lists. In addition, the valence effect was modulated by arousal level. False recall was higher

for negative/high arousal lists than for positive/high arousal lists, $p < .0001$, but did not differ between negative/low arousal lists and positive/low arousal lists.

University B

A main effect of valence was found, $F(1, 165) = 6.42$, $MSE = .05$, $p = .01$, partial $\eta^2 = .04$, as well as a Valence \times Arousal interaction, $F(1, 165) = 15.33$, $MSE = .05$, $p = .0001$, partial $\eta^2 = .09$. False recall was higher for negative lists than for positive lists, $p = .01$. Also, false recall was higher for negative/high arousal lists than for positive/high arousal lists, $p = .0005$, but did not differ between negative/low arousal lists and positive/low arousal lists. Thus, the qualitative patterns for false recall were the same for Universities A and B.

University C

There was a valence main effect, $F(1, 183) = 18.32$, $MSE = .04$, $p < .0001$, partial $\eta^2 = .09$, and a Valence \times Arousal interaction, $F(1, 183) = 9.21$, $MSE = .04$, $p = .003$, partial $\eta^2 = .05$. Negative lists elevated false recall compared to positive lists. Also, the difference in false recall was only reliable between negative/high arousal lists and positive/high arousal lists, $p < .0001$. Thus, the qualitative patterns for the valence and arousal effects on false recall were the same in subjects from all three universities.

False recognition

Aggregated results

The aggregated results for the three universities yielded a valence main effect, $F(1, 623) = 59.92$, $MSE = .06$, $p < .0001$, partial $\eta^2 = .09$, such that false recognition was higher for negative lists than for positive lists. No arousal effect or Valence \times Arousal interaction was observed. Thus, negative valence increased false recognition regardless of arousal level.

University A

The University A data produced a valence main effect, $F(1, 264) = 29.10$, $MSE = .07$, $p < .0001$, partial $\eta^2 = .10$, and a Valence \times Arousal interaction, $F(1, 264) = 6.14$, $MSE = .07$, $p = .02$, partial $\eta^2 = .02$. According to the post hoc tests, negative lists produced more false recognition than positive lists, $p < .0001$. In addition, negative/high arousal lists produced higher false recognition than positive/high arousal lists, $p < .0001$, but negative/low arousal lists and positive/low arousal lists did not differ.

University B

There was a valence main effect, $F(1, 168) = 4.20$, $MSE = .05$, $p = .04$, partial $\eta^2 = .002$, such that false recognition of critical distractors was higher for negative lists than for positive lists. However, unlike University A, there was no reliable Valence \times Arousal interaction.

University C

The results for University C were the same as those for University B. We found a valence main effect, $F(1, 183) = 26.46$, $MSE = .06$, $p < .0001$, partial $\eta^2 = .13$, such that false recognition was higher for negative lists than for positive lists, but no Valence \times Arousal interaction. Taken together, then, the data of Universities B and C provide evidence that negative valence can increase false recognition even when arousal is low.

True recall

Aggregated results

Concerning true recall, the aggregated results for the three universities produced a valence main effect, $F(1, 684) = 141.50$, $MSE = .004$, $p < .0001$, partial $\eta^2 = .17$; an arousal main effect, $F(1, 684) = 26.05$, $MSE = .004$, $p < .0001$, partial $\eta^2 = .04$; and a Valence \times Arousal interaction, $F(1, 684) = 64.49$, $MSE = .004$, $p < .0001$, partial $\eta^2 = .09$. Post hoc tests revealed that positive lists produced more true recall than negative lists, and high arousal lists produced more true recall than low arousal lists. In addition, although the difference in true recall between positive/high arousal lists and negative/high arousal lists ($p < .0001$) was much larger than the corresponding difference between positive/low arousal lists and negative/low arousal lists ($p = .03$), both differences were reliable.

University A

There was a main effect for valence, $F(1, 331) = 117.77$, $MSE = .003$, $p < .0001$, partial $\eta^2 = .26$; a main effect for arousal, $F(1, 331) = 7.30$, $MSE = .003$, $p = .007$, partial $\eta^2 = .02$; and a Valence \times Arousal interaction, $F(1, 331) = 84.56$, $MSE = .003$, $p < .0001$, partial $\eta^2 = .20$. The post hoc tests revealed that positive lists produced more true recall than negative lists, and true recall was better for high arousal lists than for low arousal lists. In addition, the positive/high arousal lists produced more true recall than the negative/high arousal lists, $p < .0001$, but the positive/low arousal lists and negative/low arousal lists did not differ.

University B

We again found a main effect for valence, $F(1, 165) = 25.03$, $MSE = .003$, $p < .0001$, partial $\eta^2 = .13$, and a main effect for arousal, $F(1, 165) = 15.72$, $MSE = .003$, $p = .0001$, partial $\eta^2 = .09$, but no interaction between them. The reasons for the main effect were the same as in University A subjects: True recall was better for positive lists than for negative lists, and was better for high arousal lists than for low arousal lists. The absence of an interaction provided the first evidence of valence effects on recall that do not depend on level of arousal.

University C

These results were the same as those for University A subjects. There were a valence main effect, $F(1, 183) = 21.25$, $MSE = .005$, $p < .0001$, partial $\eta^2 = .10$, an arousal main effect, $F(1, 183) = 7.31$, $MSE = .005$, $p = .008$, partial $\eta^2 = .04$, and a Valence \times Arousal interaction, $F(1, 183) = 7.98$, $MSE = .005$, $p = .005$, partial $\eta^2 = .04$. True recall was better for positive lists than for negative lists, and better for high arousal lists than for low arousal lists. The valence effect was moderated by the arousal, with the difference in true recall only being reliable between positive/high arousal lists and negative/high arousal lists, $p < .0001$.

True recognition

Aggregated results

When the true recognition data of the three subject samples were aggregated, the results showed that there was a valence main effect, $F(1, 623) = 76.37$, $MSE = .02$, $p < .0001$, partial $\eta^2 = .11$, and neither an arousal main effect nor an interaction between arousal and valence were observed. Thus, true recognition was better for positive lists than for negative lists, and arousal neither directly influenced true recognition nor moderated the valence effect.

University A

We found only a main effect of valence, $F(1, 264) = 9.72$, $MSE = .02$, $p = .002$, partial $\eta^2 = .04$. Similar to recall, the positive lists produced more true recognition than negative lists, but unlike recall, there was no arousal main effect, and the valence effect did not depend on the level of arousal.

University B

Similar to the results from University A, only a valence main effect was obtained, $F(1, 168) = 57.69$, $MSE = .02$, $p < .0001$, partial $\eta^2 = .26$. True recognition was again better for positive lists than for negative lists.

University C

The results from University C displayed the same pattern as those from Universities A and B. There was a valence main effect, $F(1, 183) = 17.37$, $MSE = .02$, $p < .0001$, partial $\eta^2 = .09$, but no arousal effect or interaction. True recognition was again better for positive lists than for negative lists.

Discussion

Our general aim in this article has been to advance methodological standardization across studies of how emotional content influences false memory by providing a set of emotional DRM lists (the CEL; Table 1) that have been normed on three subject samples. The valence and arousal levels of the critical distractors and list words vary systematically over the lists. This feature allows investigators to determine how these properties of emotional content influence true and false memory under a variety of theoretically motivated conditions—for instance, immediate versus delayed testing, surface versus semantic encoding, and intentional versus incidental learning, and across a number of important individual difference dimensions—for instance, age, attitude, clinical diagnoses, cognitive ability, and personality.

In that connection, the advantage of studying emotion-false memory effects with DRM-type emotional lists is that the DRM procedure has key methodological strengths. Specifically, it is able to produce robust false memory with a procedure that is both very simple and highly adaptable. Concerning simplicity, as we saw in the present norming study, only a brief induction phase, in which subjects study some short word lists, is required to produce reliable levels of false recall and false recognition. Indeed, under certain circumstances, false memories can be detected even only a few seconds after the presentation of a single four-word DRM list (Abadie and Camos 2019; Atkins and Reuter-Lorenz 2011). With respect to adaptability, the DRM procedure is so flexible that it can be adjusted to meet the requirements of virtually any type of study. Crucially, this includes the restrictive requirements of fMRI studies (Dennis et al. 2012; Kurkela and Dennis 2016) and studies of cognitively impaired populations (Brainerd et al. 2006; Budson et al. 2006).

The valence, arousal and BAS ratings for the critical distractors and list words of the CEL are reported in Table 1, with the counterbalancing of valence and arousal illustrated in Table 2. The recall and recognition data for the CEL are summarized in Tables 3 and 4, respectively. When the CEL was initially developed, the mean valence and arousal ratings for many words were missing, because they were not rated in the ANEW (Bradley and Lang 1999) norm. Accordingly, in Table 1, we have provided mean valence and arousal ratings for the CEL words using the WKB (Warriner et al. 2013)

norm instead, which contained as 13 times many words as the ANEW norms. Besides the mean valence and arousal intensity ratings, we have also included *SDs* for these ratings in Table 1. According to the emotional-ambiguity hypothesis (Brainerd 2018; Mattek et al. 2017), the effects of valence and arousal on psychological processes are moderated by uncertainty in subjective valence ratings, which can be estimated with *SDs*. However, as this hypothesis was proposed so recently, prior emotion-memory studies have not included the ambiguity factor. The variability data in the CEL norms allow investigators either to manipulate ambiguity factorially or to include it as a covariate when either valence or arousal is manipulated.

By far the most important feature of the norming results is the level of consistency in the effects that were observed for the three subject samples. A major criticism of the extant emotion-false memory literature is the inconsistent findings for valence and arousal manipulations across experiments (see Bookbinder and Brainerd 2016). However, in the current study, the analyses of the false recall, false recognition, true recall, and true recognition data for the individual universities produced the same qualitative patterns, with only two minor exceptions. Those exceptions involved the false recognition data and the true recall data. In false recognition, a reliable Valence \times Arousal interaction was obtained for University A but not for University B, University C or the aggregated data. In true recall, a reliable Valence \times Arousal interaction was obtained for University A, University C, and the aggregated data but not for University B. Considering the number of ANOVAs that were conducted, it is not surprising, from a statistical point of view, that two minor findings varied among the subject samples.

Overall, false recall and false recognition were more marked for negative lists than for positive lists (see Tables 3 and 4). Neither type of false memory varied consistently as a function of arousal. These results echo Brainerd and Bookbinder's (2019) finding that compared to arousal, valence is more strongly correlated with semantic properties that reliably predict false memory (e.g., meaningfulness, familiarity). Also, the effects of valence on false memory were moderated by arousal for false recall but not false recognition: The difference between negative versus positive valence in false recall widened as arousal increased. Notably, valence affected true memory differently than it affected false memory. In particular, true recall and true recognition were better with positive lists than with negative lists, which is consistent with other studies that have factorially manipulated valence and arousal (Brainerd et al. 2010; Gomes et al. 2013; Libkuman et al. 2004). Empirically, this is a classic example of a double dissociation, wherein a manipulation has opposite effects on different performance measures. The specific pattern of opposite valence effects for true and false memory is predicted by fuzz-trace theory (FTT)'s account of emotion-false memory

effects. According to that account (see Bookbinder and Brainerd 2016; Brainerd and Bookbinder 2019), negative content enhances gist memory (memory for items' semantic content and other relational information, which is the primary process that underlies false recall and false recognition), but positive content enhances verbatim memory (memory for item-specific surface details, which is the primary process that underlies true recall and true recognition).

The fact that for both recall and recognition, negative lists produced more false memory than positive lists is consistent with some prior studies that used DRM-type materials (Brainerd et al. 2008a; Howe 2007; Howe et al. 2010; Sharkawy et al. 2008). These results have important implications in law and other real-life settings, where it is often claimed that memories of negative events are especially well preserved and highly resistant to distortion (Laney and Loftus 2010). It is worth mentioning that in most of the prior emotional DRM studies, valence was confounded with arousal, which we managed to avoid in the current study. Importantly, we observed a robust interaction between valence and arousal for false recall, but not for false recognition: The valence effect on false recall was much larger when arousal was high than when it was low. Based on Bookbinder and Brainerd's (2016) review, that result seems to be new to the CEL. Why does arousal modulate the valence effect in false recall but not in false recognition? One possibility lies in the difference between recall and recognition tests that the former is more sensitive to verbatim memory, whereas the latter is more sensitive to gist memory (Seamon et al. 2003). As FTT predicts, the arousal effect is predominantly rooted in verbatim memory (Brainerd and Bookbinder 2019). Specifically, verbatim memory should improve as arousal increases from low to moderate but decline as arousal increases from moderate to high (Bookbinder 2017). Thus, owing to the fact that recall tests favor verbatim memory more than recognition tests, it is expected that arousal should have a stronger effect on recall than on recognition.

Continuing with the detailed norming results, we found for both recall and recognition that positive content elevated true memory relative to negative content. In addition, arousal moderated the valence effect in recall but not in recognition, just as it did in false memory, and it also had a main effect. High arousal lists produced more true recall than low arousal lists, and there was an interaction such that the valence effect was more marked for high arousal lists. Again, the arousal modulation effect was only found for recall but not recognition, and the difference in sensitivity to verbatim memory between recall and recognition tests may be responsible. If the arousal effect is primarily verbatim-based, as FTT expects, it would be more likely to show up in recall tests than in recognition tests.

It should be noted that the differences between recall and recognition that were detected in the CEL norming data are not without precedent in the emotion-false memory literature.

Two points that were emphasized in Bookbinder and Brainerd's (2016) review are that across experiments in this literature, (a) recall experiments have produced different emotion-false memory effects than recognition experiments, and (b) recognition effects have been more consistent than recall effects. However, as these authors also noted, it is difficult to conclude anything more than that because variations in valence were confounded with variations in arousal in all studies. That removal of that persistent confound is one of the great advantages of the CEL. Based on our norming results, recall and recognition do indeed produce different emotion-false memory effects with DRM-type tasks, and they also seem to produce different emotion-true memory effects.

Summing up, the current study provides norming data for 32 lists of which all list words and critical distractors have been normed for perceived valence (positive-negative) and perceived arousal (calm-exciting). These lists have been normed for the levels of true recall, true recognition, false recall, and false recognition that they generate. The most robust patterns are that true memory increased and false memory decreased when the content of list words or critical distractors were positive compared to when they were negative. These patterns hold for both recall and recognition in the norming data, for all subject samples. Another consistent pattern is the Valence \times Arousal interactions in recall, such that high arousal always amplifies the effects of valence. Because the valence and arousal of target materials have been so routinely confounded in prior research, a key point about this pattern is that it emerged from factorial manipulations of valence and arousal, which the CEL makes possible.

References

- Abadie, M., & Camos, V. (2019). False memory at short and long term. *Journal of Experimental Psychology: General*, *148*(8), 1312–1334.
- Atkins, A. S., & Reuter-Lorenz, P. A. (2011). Neural mechanisms of semantic interference and false recognition in short-term memory. *NeuroImage*, *56*(3), 1726–1734.
- Bookbinder, S. H. (2017). *Effect of low- and high- arousal mood on true and false memory for neutral words*. (Unpublished doctoral dissertation). Cornell University, NY, USA.
- Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context-content paradox. *Psychological Bulletin*, *142*(12), 1315–1351.
- Bookbinder, S. H., & Brainerd, C. J. (2017). Emotionally negative pictures enhance gist memory. *Emotion*, *17*(1), 102–119.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25–36). Technical report C-1, the center for research in psychophysiology, University of Florida.
- Brainerd, C. J. (2018). The emotional ambiguity hypothesis: A large-scale test. *Psychological Science*, *29*(10), 1706–1715.
- Brainerd, C. J., & Bookbinder, S. H. (2019). The semantics of emotion in false memory. *Emotion*, *19*(1), 146–159.
- Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford University Press.
- Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology*, *42*(5), 662–679.
- Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008a). How does negative emotion cause false memories?. *Psychological Science*, *19*(9), 919–925.
- Brainerd, C. J., Yang, Y., Togliani, M. P., Reyna, V. F., & Stahl, C. (2008b). *Emotion and false memory: The Cornell/Cortland lists*. In *annual meeting of the Psychonomic Society*, Chicago, IL.
- Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Togliani, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology*, *107*(2), 137–154.
- Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*(1), 419–439.
- Brueckner, K., & Moritz, S. (2009). Emotional valence and semantic relatedness differentially influence false recognition in mild cognitive impairment, Alzheimer's disease, and healthy elderly. *Journal of the International Neuropsychological Society*, *15*(2), 268–276.
- Budson, A. E., Todman, R. W., Chong, H., Adams, E. H., Kensinger, E. A., Krangel, T. S., & Wright, C. I. (2006). False recognition of emotional word lists in aging and Alzheimer disease. *Cognitive and Behavioral Neurology*, *19*(2), 71–78.
- Choi, H. Y., Kensinger, E. A., & Rajaram, S. (2013). Emotional content enhances true but not false memory for categorized stimuli. *Memory & Cognition*, *41*(3), 403–415.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17–22.
- Dennis, N. A., Bowman, C. R., & Vandekar, S. N. (2012). True and phantom recollection: An fMRI investigation of similar and distinct neural correlates and connectivity. *NeuroImage*, *59*(3), 2982–2993.
- Gallo, D. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. Psychology Press.
- Gallo, D. A., Foster, K. T., & Johnson, E. L. (2009). Elevated false recollection of emotional pictures in young and older adults. *Psychology and Aging*, *24*(4), 981–988.
- Gomes, C. F., Brainerd, C. J., & Stein, L. M. (2013). Effects of emotional valence and arousal on recollective and nonrecollective recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 663–677.
- Howe, M. L. (2007). Children's emotional false memories. *Psychological Science*, *18*(10), 856–860.
- Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory*, *18*(1), 58–75.
- Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory & Language*, *52*(1), 1–28.
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior*, *34*(1), 3–38.
- Kurkela, K. A., & Dennis, N. A. (2016). Event-related fMRI studies of false memory: An activation likelihood estimation meta-analysis. *Neuropsychologia*, *81*, 149–167.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). <https://doi.org/10.18637/jss.v082.i13>
- Laney, C., & Loftus, E. F. (2010). Truth in emotional memories. In B. H. Bornstein & R. L. Wiener (Eds.), *Emotion and the law* (pp. 157–183). Springer.
- Libkuman, T., Stabler, C., & Otani, H. (2004). Arousal, valence, and memory for detail. *Memory*, *12*(2), 237–247.

- Lindsay, D. S., & Read, J. D. (1994). Psychotherapy and memories of childhood sexual abuse: A cognitive perspective. *Applied Cognitive Psychology, 8*(4), 281–338.
- Lynn, S. J., Lock, T., Loftus, E. F., Krackow, E., & Lilienfeld, S. O. (2003). The remembrance of things past: Problematic memory recovery techniques in psychotherapy. In S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 205–239). Guilford Press.
- Mattek, A. M., Wolford, G. L., & Whalen, P. J. (2017). A mathematical model captures the structure of subjective affect. *Perspectives on Psychological Science, 12*(3), 508–526.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers, 36*(3), 402–407.
- Palmer, J. E., & Dodson, C. S. (2009). Investigating the mechanisms fuelling reduced false recall of emotional material. *Cognition and Emotion, 23*(2), 238–259.
- Pesta, B. J., Murphy, M. D., & Sanders, R. E. (2001). Are emotionally charged lures immune to false memory?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 328–338.
- Poole, D. A., & Lamb, M. E. (1998). *Investigative interviews of children*. Washington, DC: American Psychological Association.
- Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*(3), 715–734.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803–814.
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*(3), 385–407.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178.
- Seamon, J. G., Goodkind, M. S., Dumey, A. D., Dick, E., Aufseeser, M. S., Strickland, S. E., Woulfin, J. R., Fung, N. S. (2003). “If I didn’t write it, why would I remember it?” Effects of encoding, attention, and practice on accurate and false memory. *Memory & Cognition, 31*(3), 445–457.
- Sharkawy, J. E., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False memories of emotional and neutral words. *Behavioural Neurology, 19*(1–2), 7–11.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50.
- Stadler, M. A., Roediger, H. L., III, & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition, 27*(3), 494–500.
- Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207.
- Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest, 18*(1), 10–65.
- Wundt, W. M. (1912). *An introduction to psychology* (R. Pintner, Trans. & M. D’Hoogh Ed.). Allen & Unwin.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.