

# Normative Emotional Responses to Behavior Analysis Jargon or How Not to Use Words to Win Friends and Influence People

Thomas S. Critchfield<sup>1</sup> · Karla J. Doepke<sup>1</sup> · L. Kimberly Epting<sup>2</sup> · Amel Becirevic<sup>3</sup> · Derek D. Reed<sup>3</sup> · Daniel M. Fienup<sup>4</sup> · Jamie L. Kreamsreiter<sup>5</sup> · Cheryl L. Ecott<sup>6</sup>

Published online: 27 February 2017  
© Association for Behavior Analysis International 2016

**Abstract** It has been suggested that non-experts regard the jargon of behavior analysis as abrasive, harsh, and unpleasant. If this is true, excessive reliance on jargon could interfere with the dissemination of effective services. To address this often discussed but rarely studied issue, we consulted a large, public domain list of English words that have been rated by members of the general public for the emotional reactions they evoke. Selected words that behavior analysts use as technical terms were compared to selected words that are commonly used to discuss general science, general clinical work, and behavioral assessment. There was a tendency for behavior analysis terms to register as more unpleasant than other kinds of professional terms and also as more unpleasant than English words generally. We suggest possible reasons for this finding, discuss its relevance to the challenge of deciding how to communicate with consumers who do not yet understand or value behavior analysis, and advocate for systematic research to guide the marketing of behavior analysis.

**Keywords** Jargon · Behavior analysis terminology · Dissemination · Motivating operations · Emotion

✉ Thomas S. Critchfield  
tsritic@ilstu.edu

- <sup>1</sup> Department of Psychology, Illinois State University, Normal, IL 61761, USA
- <sup>2</sup> Elon University, Elon, NC, USA
- <sup>3</sup> University of Kansas, Lawrence, KS, USA
- <sup>4</sup> Queens College, Flushing, NY, USA
- <sup>5</sup> University of Iowa Hospitals and Clinics, Iowa City, IA, USA
- <sup>6</sup> Better Life Behavioral Services of Central Florida, Leesburg, FL, USA

If published commentaries are of any guide, applied behavior analysts tend to agonize over how they are perceived by laypersons and other non-experts (e.g., Bailey, 1991; Doughty et al., 2012; Freedman, 2015; Foxx, 1996; Lindsley, 1991; Smith, 2015)—and possibly for good reason. Maurice (1993), for instance, wrote that in her experience, a behavior analyst often is seen as “Attila the Hun”<sup>1</sup> rather than the “angel of love and acceptance” (p. 283) that laypersons expect of service professionals. The core concern, therefore, involves not competence so much as social acceptability.

Three general observations apply to this concern. First, it suggests a significant roadblock to fulfilling the ethical-social responsibility of getting effective behavior analysis interventions to people who need them. Consumers tend to seek out and follow the suggestions of therapists with whom they feel socially comfortable (Backer et al., 1986; Barrett-Lennard, 1962; Rosenzweig, 1936). Unfavorable public perception thus creates a challenge of marketing and implementation (Bailey, 1991; Doughty et al., 2012; Lindsley, 1991).<sup>2</sup>

Second, frequently implicated in behavior analysts’ reflections on their own poor public relations is a peculiar way of speaking that non-behavior analysts are thought to regard as “abrasive” (Lindsley, 1991, p. 449), “harsh” (Maurice, 1993, p. 102), and generally unpleasant (Bailey, 1991). The underlying problem, as Foxx (1996) saw it, is that behavior analysis terms often repurpose words that have unpleasant everyday connotations (see also Lindsley, 1991). To illustrate, Foxx (1990) provided what he called “North American translations” of a number of behavior analysis terms, including

Chaining—foreplay by bondage devotees.

<sup>1</sup> For whatever it is worth, the historical record portrays Attila as ruthlessly efficient but, at unpredictable intervals, surprisingly compassionate and ethical (Howarth, 1994).

<sup>2</sup> In which we perceive no small irony given that a behaviorist, John B. Watson, is credited as a major innovator in modern marketing (Buckley, 1982).

Discrimination—prevented by 1964 federal law.

Extinction—the disappearance of a species. Can cause concern for parents when they are told, “We could like to put your son, Jason, on extinction” (p. 950).

Because of pre-existing associations, non-experts may experience behavior analytic terms very differently from how behavior analysts experience them. Whereas behavior analysts use technical terms in an attempt to promote clear understanding of behavioral processes, they may elicit unpleasant, gut-level emotional reactions instead. Regarding implications of this for applied behavior analysis, Foxx (1996) reminded that, “People’s emotional reactions are critical to successful program adoption and that behaviorally induced resistance to change can sabotage any program” (p. 157).<sup>3</sup>

Third, the notion that behavior analysis terms have abrasive emotional overtones may be intuitively appealing but, as far as we are aware, not empirically verified. When behavior analysts have discussed the role of language in their own public relations, anecdote usually has substituted for systematic evidence. Only research can verify whether a hypothesized communication gap is real and, if so, what its parameters might be (for suggestive evidence, see Becirevic et al., 2016; Witt et al., 1984). Here, we describe a simple means of gaining a preliminary empirical understanding of how members of the general public might react to some of the words that comprise behavior analysis terminology.

## Method

### Source of Data

Our data source was a large, public domain list of nearly 14,000 English words that have been rated for how they strike people emotionally (Warriner et al., 2013). Hereafter, we refer to this as the Warriner corpus. In service of brevity, here we describe only selected aspects of the rating procedure; for additional information, see the “Appendix” and Warriner et al. (2013).

The words were rated by volunteers in Amazon Mechanical Turk (mTurk), an online data collection platform. Previous research suggests that volunteers in large-scale mTurk studies tend to approximate US population demographics better than traditional academic setting convenience samples (Paolacci & Chandler, 2014). It is in this sense that the Warriner corpus ratings serve as a normative estimate of how typical non-experts might be expected to respond.<sup>4</sup>

<sup>3</sup> One relevant finding is that positively emotional communication makes a speaker seem more familiar (Garcia-Marques et al., 2004). If the opposite is true—that negatively emotional communication makes speakers seem strange or remote—then those who use unpleasant words are unlikely to be the “comfortable” therapists that consumers prefer (Backer et al., 1986; Barrett-Lennard, 1962; Rosenzweig, 1936).

<sup>4</sup> In Warriner et al. (2013), the raters were about 60% female and reflected a broad range of ages and levels of education.

The emotion ratings of the Warriner corpus are a sample of listener behavior. mTurk volunteers rated their own emotional responses to each word on three dimensions, two of which are relevant here. The first dimension was a scale of 1 (“unhappy”) to 9 (“happy”). In the literature on emotional responding to words (e.g., Bradley & Lang, 1999), this is a fairly standard way of characterizing emotional valence. “Unhappy” and “happy” anchors serve as a user-friendly means of capturing the extent to which a given word elicits general unpleasant emotions, in the sense of *this word makes me feel unhappy* (or “annoyed, unsatisfied, melancholic, despaired, or bored”; Warriner et al., 2013, p. 1193) versus general pleasant emotions, in the sense of *this word makes me feel happy* (or “pleased, satisfied, contented, hopeful”; p. 1193). Consequently, in the **Results and Discussion** section, we describe ratings as ranging from “unpleasant” to “pleasant.”

The second dimension, arousal, was a scale of 1 (“calm”) to 9 (“excited”). Arousal implies something like strength of behavioral activation, about which a word of conceptual framing is in order. Skinner (1953) described emotions in terms of motivating operations or conditions that make it more or less reinforcing to engage in particular behaviors. Arousal may be thought of as addressing this feature of emotion, and consequently, in the **Results and Discussion** section, we will describe the arousal scale as ranging from “not motivating” to “motivating.”

One further note on the emotion rating procedure: mTurk volunteers viewed and rated each word separately rather than in narrative context. This approach is in keeping with the observation that listeners can respond to words independently on many levels including semantically, phonologically, and orthographically (e.g., Perfetti et al., 1988). The isolated word rating procedure is an attempt to get at these sorts of responses in more or less pure form, that is, free of the influence of responses to surrounding text.

### Selection of Terms and Assignment of Emotion Ratings

We examined the entire Warriner corpus for words that we recognized as important in behavior analysis technical discussions. As a hedge against inadvertent omissions, we consulted Cooper et al. (2007) for additional terms that we had not noticed in the corpus and then double-checked the corpus for them. When a term exactly matched a word in the Warriner corpus, we assigned the valence and arousal ratings listed in the corpus for that word. Some terms of obvious potential interest (e.g., *generalization* and *efficacy*) were not part of the Warriner corpus and thus could not be included in our analysis.

Terms for which emotion ratings were available were divided into four broad categories:

*Behavior analysis technical terms* ( $N = 39$ ) that have unique meanings in discussions between expert behavior

analysts about behavioral functional relations and/or interventions

*General science terms* ( $N = 42$ ) that behavior analysts often employ but whose usage is not idiosyncratic to behavior analysis

*Behavioral assessment terms* ( $N = 34$ ) related to measuring behavior and detecting clinical effects

*General clinical terms* ( $N = 35$ ) that behavior analysts employ often but whose usage is not idiosyncratic to behavior analysis.

Note that assignment of terms to categories was based on our “expert” judgment; readers are invited to draw independent conclusions about the validity of our decisions (see Figs. 2 and 3 for terms).

## Results and Discussion

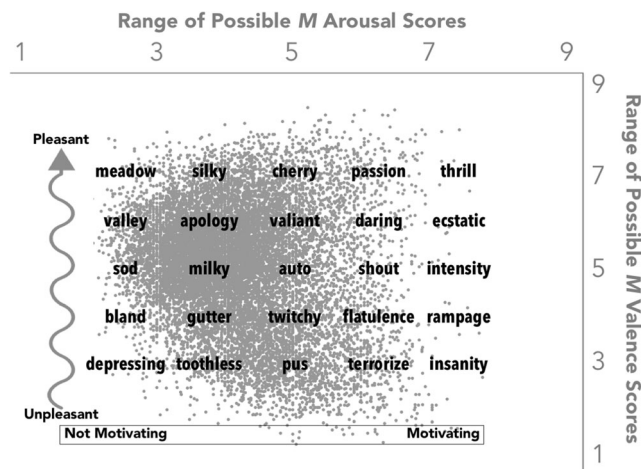
Figure 1 explains our mode of display for word-emotion ratings and puts the ratings into context by showing those of the nearly 14,000 words of the Warriner corpus (gray data points). In scatter plot format, valence ratings are scaled on the ordinate and arousal ratings are scaled on the abscissa; specific rating-scale values are shown at top and at right. The scatter plot shows that, overall, the English language appears to contain more pleasant than unpleasant words and more not-

motivating than motivating words (Dodds et al., 2015; Warriner & Kuperman, 2015).

Superimposed on the Warriner corpus scatter plot are the axes (bottom and left) used in subsequent figures. Although these axes are proportioned with numerical accuracy, numerical labels are omitted (a) to reduce visual clutter and (b) because ratings on an ordinate scale communicate relative, not precise values, and we do not wish to imply that a specific ordinal rating corresponds to a specific amount of something in the way that response counts do. Also, superimposed are some familiar everyday words (plotted like data points, i.e., centered on approximate rating scale values) that were selected to illustrate specific combinations of ratings. For example, the ratings portray both *thrill* and *meadow* as strongly pleasant, but *thrill* is more motivating. *Terrorize* and *toothless* both are strongly unpleasant, but *terrorize* is more motivating. *Ecstatic* and *rampage* are about equally motivating, but with different valences. *Auto* is unremarkable in terms of both valence and arousal.

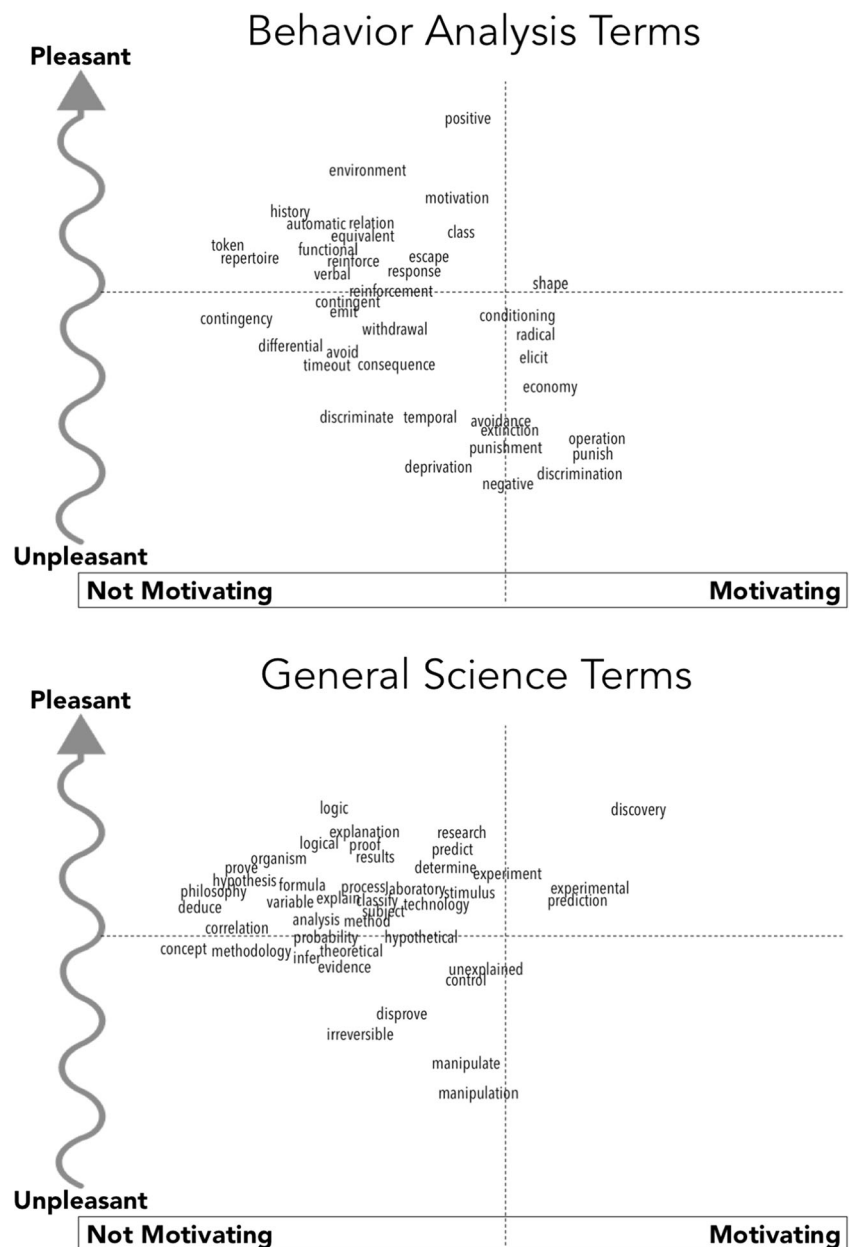
Using the axis format illustrated in Fig. 1, Fig. 2 shows the ratings for behavior analysis (top panel) and general science (bottom panel) terms, with words plotted instead of data points. Figure 3 shows the ratings for behavioral assessment (top panel) and general clinical (bottom panel) terms. In both cases, in lieu of quantitative axis labels, dashed lines indicate the middle (neutral) rating value on each scale, in effect creating a 2 (pleasant vs. unpleasant)  $\times$  2 (motivating vs. not motivating) matrix. Figure 4 summarizes by showing the percentage of words from each category that fell into each quadrant of the matrix. Also shown in Fig. 4 is the percentage of total words in the Warriner corpus that fell into each quadrant (dashed line in each panel).

Consistent with underlying distributions for English words (Fig. 1), a majority of words were rated as pleasant in the general science (67%), behavioral assessment (67%), and general clinical (53%) categories. By contrast, a majority of behavior analysis terms (60%) were rated as unpleasant. It has been suggested that, at a molar level, the emotional tone of communication is affected mainly by the most strongly valenced words (e.g., Dodds & Danforth, 2009). For example, when evaluating large samples of printed text, Reagan et al. (2015) focused mainly on words that were rated below 4 and above 6 on a nine-point “happy” scale. Applying the same heuristic to our lists of terms shows that strongly positive words outnumbered strongly negative words in the general clinical (12:9), general science (9:4), and behavioral assessment (7:5) categories. The opposite was true for behavior analysis terms (1:15). Moreover, only among behavior analysis terms was a substantial percentage of words (28%) rated as strongly motivating. Most of the relevant words had a negative valence, which is important because arousal is thought to magnify the impact of valence (Warriner & Kuperman, 2015). Overall, the available evidence is consistent with what would be expected if behavior analytic communication indeed tends to be abrasive.



**Fig. 1** Ratings of 13,915 English words in the Warriner corpus (gray data points) indexed to the valence (top axis) and arousal (right axis) word-emotion rating scales. Also shown: Some familiar words representing various combinations of valence (pleasantness) and arousal (motivation) ratings. The simplified axes (left and bottom) used here and in Figs. 2 and 3 indicate the range of ratings that they represent. Note that the ranges of values on these axes were selected to encompass the words of interest rather than the full range of possible ratings. See text for further explanation

**Fig. 2** Emotion ratings for some behavior analysis and general science terms



### Three Limitations

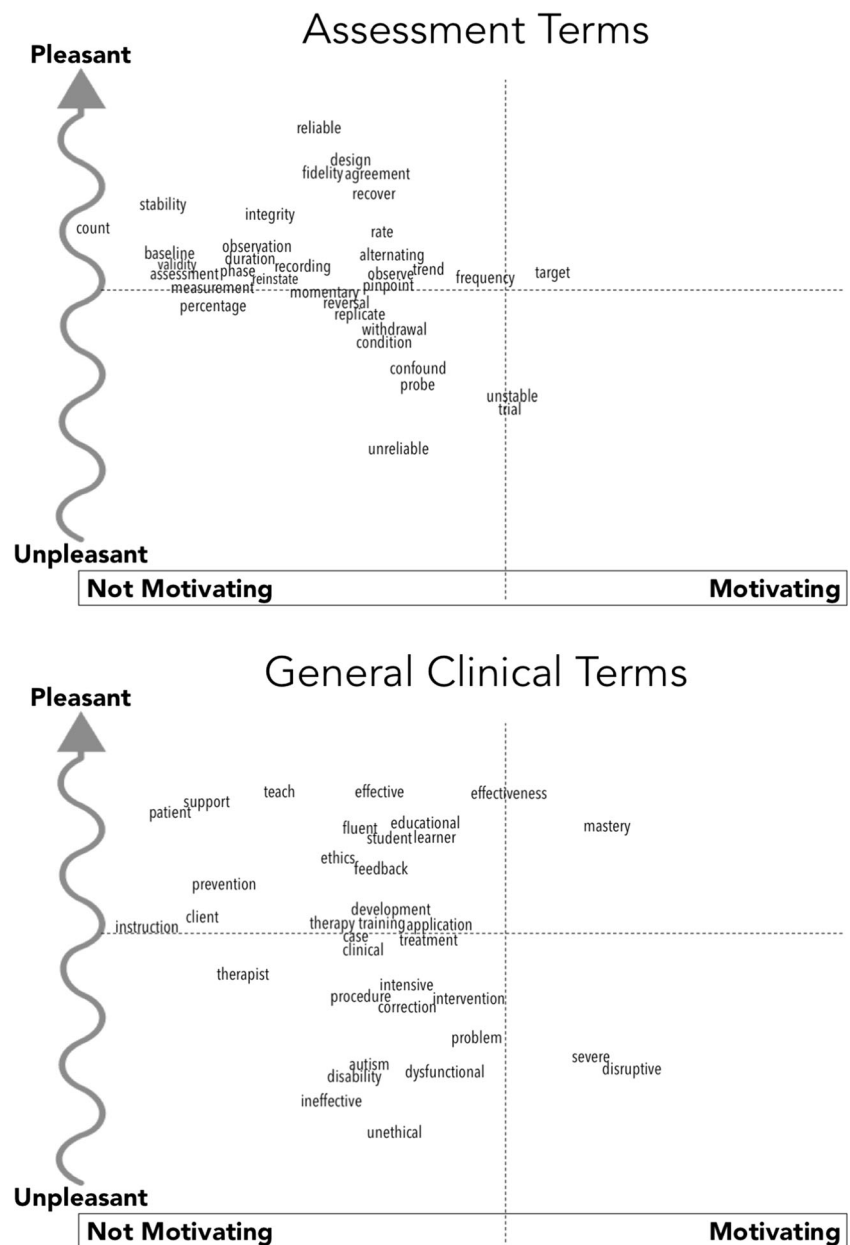
One limitation of the present study is that self-reports of listener emotional responding are *self-reports*, which behavior analysts tend not to trust as data (Baer et al., 1968). Regarding this concern we note that word-emotion ratings are a type of social validity assessment, and verbal reports typically are accepted as a convenient and informative source of social validity information (Wolf, 1978). Moreover, research from outside of behavior analysis indicates that self-reported emotional responding predicts a number of outcomes that could be relevant to dissemination. These include demand for consumer products (Floh et al., 2013), the degree to which followers trust leaders (Norman et al., 2010),

and how favorably persuasive messages are perceived (particularly when the topic at hand requires detailed consideration, as might be true of the typical clinical case; Petty et al., 1993).<sup>5</sup>

Another limitation is that our analyses were constrained by the terms that appeared in the Warriner corpus, and there is no telling how representative these might be of the overall behavior analytic lexicon. This defines a promising direction for future research. Using the easily replicated Warriner et al. (2013) word-rating

<sup>5</sup> Non-practitioners take note: There is even some evidence that the linguistic style of scientific abstracts affects the probability that an article will “go viral” via citations and professional social media communication (Guerini et al., 2012).

**Fig. 3** Emotion ratings for some assessment and general clinical terms



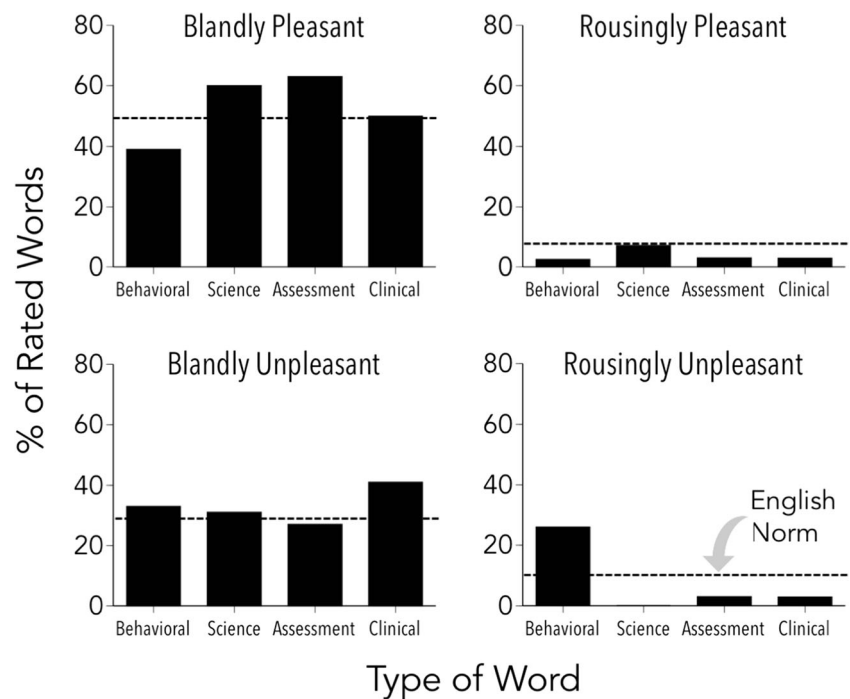
procedure, it would be useful to flesh out the range of terms for which normative listener reactions are available.

A third limitation is that, although word-emotion ratings suggest that abrasive technical terms could cause non-experts to avoid behavior analysts or reject their offers of professional assistance (Bailey, 1991; Foxx, 1996; Lindsley, 1991), the ratings do not verify conditions under which this may actually occur. Results of a few studies align with our current findings by showing that behavior analytic interventions were judged to be less desirable when described with technical jargon rather than in everyday language (Becirevic et al., 2016; Jarmolowicz et al., 2008; Witt et al., 1984), but more research of this type is needed.

### Functionally Abrasive Repurposed Terms

In the meantime, our results offer clues about the nature of the problem. A reasonable response to marketing concerns in behavior analysis is to question the user-friendliness of *all* professional communication—perhaps laypersons find the communication of expert scientists and clinicians to be generally abrasive, and behavior analysis jargon is not remarkable in this regard. Based on word-emotion ratings, however, most of the general science, behavioral assessment, and general clinical terms we examined might be described as mundane (neither invigoratingly joyful nor abrasive). Behavior analysis terms, by contrast, tended to be more unpleasant overall, and more *arousingly* unpleasant in many cases.

**Fig. 4** Percentages of terms in four categories defined by ratings of pleasant vs. unpleasant and motivating (or rousing) vs. not motivating (or bland). *Dashed lines* show percentages for the English language overall as represented by the Warriner corpus



Foxx (1996) suggested that words which behavior analysts have repurposed for their own technical uses often involve verbal topographies that carry unpleasant connotations from everyday communication (Fig. 2). This is plausible speculation, but one hopes to do better than to speculate. To explore the kinds of associations that behavior analysis technical terms might evoke (i.e., the relational networks in which they participate), we consulted two publicly available corpora of word associations that were created by asking large numbers of individuals to indicate the first word that occurred to them upon hearing a target word. For several thousand words, the corpora list these associates and their frequencies of occurrence. Unfortunately, most behavior analysis technical terms do not appear in these corpora, but for six terms that do, Fig. 5 shows some common word associates. In each panel of Fig. 5, a technical term appears in large gray letters and its common associates appear in black.<sup>6</sup> In keeping with the format of Figs. 2 and 3, word positions are determined by emotion ratings in the Warriner corpus. Font size for associates is proportional to their frequency of occurrence (larger = more frequent, i.e., stronger normative association with the target word). We

could not plot associates that are not rated in the Warriner corpus, and to reduce visual clutter, we omitted some low-frequency associates. These limitations notwithstanding, Fig. 5 aligns nicely with the claim that some behavior analysis terms evoke unpleasant associations (Foxx, 1996). Such associations may be thought of as a measure of the verbal history that must be overcome when introducing a typical non-expert to behavior analysis terms.

As a convenient mnemonic and nod to Foxx's (1996) conceptual analysis of the communication problem, we will refer to verbal topographies that elicit unpleasant emotional responses, presumably due to listener experience in the lay verbal community, as *functionally abrasive repurposed terms* (FRTs). Based on the present evidence, our preliminary advice to practitioners, like that of previous writers (e.g., Foxx, 1990; Lindsley, 1991), is to employ care in communicating with non-experts. This advice especially concerns the role of FRTs early in a therapist-consumer relationship. In professional interactions (as in all social relationships), first impressions matter (Corrigan et al., 1980). For example, when consumer and therapist come together for the first time, they are strangers (mutually neutral stimuli). Once the therapist has become a conditioned reinforcer, through pairing with constructive conversations and satisfying therapeutic outcomes (for examples see Maurice, 1993), the therapist-consumer relationship is resistant to perturbation, such that an accidentally offensive comment may not send the consumer dashing away. When the therapist is still a neutral stimulus, however, pairing with even a mildly aversive event (as when the therapist emits a FRT) can substantially influence the trajectory of the relationship.

<sup>6</sup> Three terms (*punish*, *shape*, and *escape*) were evaluated using the University of South Florida Free Association Norms (<http://www.usf.edu/FreeAssociation/>). This corpus provides the proportion of individuals who generated a forward (word to associate) match and the proportion who generated a backward (associate to word) match; Fig. 5 reflects the sum of these. Three other terms (*operation*, *avoid*, and *extinct[ion]*) were evaluated using the Edinburgh Associative Thesaurus (<http://www.eat.rl.ac.uk/>). This corpus provides only forward association data. Thus, although the source and format of data varied across panels, each panel provides on means of estimating high-probability word associations.

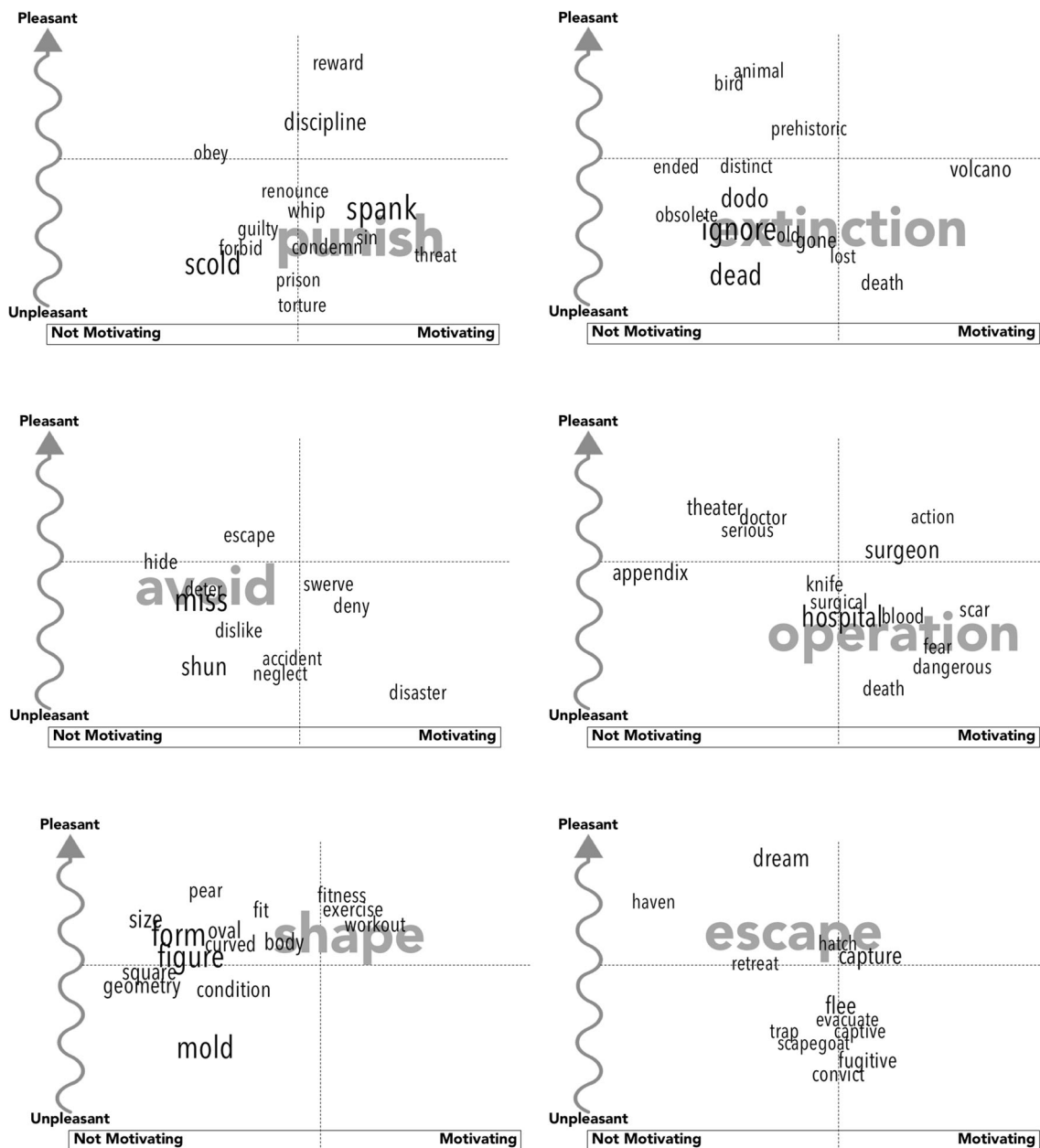


Fig. 5 Common word associates of six example terms. See text for details

Too often, we submit, FRTs cast an unproductive haze over the interactions between behavior analysts and the non-experts they hope to recruit as clients (or students, or colleagues, or professional allies). Berger (1973) has detailed how, in an interdisciplinary setting, when behavior analysts began emitting FRTs such as “conjunctive schedules and respondent discrimination” (p. 106), their non-behavioral colleagues would “turn pale and begin to leave the room” (p. 106). This anecdote, extended metaphorically, suggests a useful maxim—*People who emit too many FRTs find themselves alone in the room*—in view of which, we suggest that solitude is a poor vantage point from which to accomplish good in the everyday world.

### Diffusing Feedback Deafness

The core concern addressed in the present report is that non-experts may neither like nor feel informed by behavior analytic jargon. Anecdotal accounts suggest that they have been signaling this for a long time (see Berger, 1973; Lindsley, 1991). It is therefore interesting that the vocabulary of behavior analysis has remained remarkably unchanged since it was first coined (primarily in the 1930s through the 1950s), even as attention shifted partly from discovery toward application and dissemination (e.g., Baer, et al., 1968). In considering the needs of new audiences encountered during this process, Foxx (1996) proposed that dissemination has been impaired by

“feedback deafness” (p. 150), in which behavior analysts stubbornly adhere to verbal practices that arose in service of theoretical, rather than practical, considerations. Skinner himself could be guilty of inflexibility, as when he opined that, “the skepticism of [non-behavior analysts] about the adequacy of behaviorism is an inverse function of the extent to which they understand it” (1988, p. 472). Yet, verbal behavior is a functional phenomenon in which there is no “meaning” except in the speaker-listener relations that actually emerge.

To state the matter more simply, how listeners are affected by verbal behavior has no necessary connection to what speakers believe they are saying. A robust literature on dissemination (e.g., Rogers, 2003) indicates that blind hope to the contrary cannot substitute for the laborious process of “conducting a front-end analysis with potential consumers to discover exactly what they were looking for, what form it should take, and *how it should be packaged and delivered*” (Bailey, 1991, p. 446, italics added). Concerns about behavior analysis jargon thus mirror a growing contemporary “healthcare literacy” movement (e.g., Coleman, 2011), in which choice of phrasing is viewed as an important “packaging” consideration. Specifically, the movement focuses on understanding how language functions differently for patients than it does for healthcare providers and educators and seeks empirical guidance regarding the best ways to communicate with non-experts. Among the goals are the following: when necessary to help non-experts master technical information; when possible to find non-technical ways of communicating; and always to employ objective means of determining how to meet these objectives.

By no means are the present data intended as a final word on the social validity of behavior analysis terms. There are many ways to empirically evaluate the functional properties of terms, and using publicly available data sets of emotion ratings to predict normative listener reactions is merely a convenient place to start. More important than any specific finding of the present study is its guiding value: that listener reactions will be better understood through research than through the qualitative impressions that have guided much previous commentary on this issue.

A shortage of empirical guidance means that it is premature to offer evidence-based “best practices” for communicating with non-experts. Instead, we advance three general considerations for those who are concerned with this issue. First, interactions with non-experts should be thought of, not as “conveying information,” but rather as speaker-listener relations in which “the organism (listener) is always right.” Second, emotion elicitation is but one function of verbal responses; also of interest to practitioners are additional functions such as how well various kinds of instructions control listener behaviors involved in therapeutically essential tasks like implementing interventions and collecting progress data (Jarmolowicz et al. 2008). Third, listeners have varied learning histories, and so “best practices” for communicating may

vary across audiences. The present study may be said to have treated “non-experts” as a generic population, but this is a convenient fiction. Verbal repertoires differ across individuals and also across groups that are defined by such experiences as immersion in the majority language (e.g., in the USA, native English speakers versus speakers of English as a second language), professional training (e.g., nurses versus social workers versus special education teachers), and previous consultation history (e.g., parents of children who are newly diagnosed versus veterans of extensive prior treatment).

As our field awaits the research that can guide better communication, an unfortunate reality to be acknowledged is that subjective guesses about how to connect with listener repertoires can be fallible. Consider the case of Ogden Lindsley, a pioneer advocate of matching communication style to consumer needs. Lindsley (1991, Table 1) distilled “25 years translating ... technical jargon into plain English for use by public school children, parents, and teachers” (p. 449) into a list of terms that he believed communicated effectively and avoided the abrasiveness of technical jargon. Using the Warriner corpus as a metric, some of Lindsley’s preferred terms indeed qualify as pleasant (e.g., on the nine-point “happy” scale, *precision* = 6.38; *reward* = 7.47; [*ac*] *celeration* = 6.19), but others do not (e.g., *consequence* = 3.86; *penalty* = 2.8; *manager* = 4.82). Apparently “25 years translating” was insufficient for Lindsley to fully anticipate the gut-level emotional responses that words can evoke. This illustrates why behavior analysts, who excel in investigating behavior-environment relations, need a thoroughgoing empirical analysis of the role of their own verbal practices as part of the environment in which the behavior of laypersons and other non-experts unfolds.

### Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Ethical Approval** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

This article does not contain any studies with human participants performed by any of the authors. It drew instead on archival, public domain data sets. Therefore, Institutional Review Board oversight does not apply.

**Informed Consent** Because the research used archival, public domain data sets, conventions of informed consent do not apply.

## Appendix

### Summary of the Warriner et al. (2013) Rating Procedure

Each participant was paid 75 US cents for rating a set of approximately 350 words. A rater evaluated a given set of



words on only one emotional dimension (e.g., only pleasantness or arousal). The first 10 words provided raters with practice using the full range of possible ratings and were drawn from lists of words shown in previous research to evoke a wide range of emotional responses (for pleasantness: jail, invader, insecure, industry, icebox, hat, grin, kitten, joke, and free; for arousal: statue, rock, sad, cat, curious, robber, shotgun, assault, thrill, and sex). Of the remaining words, 40 were drawn from a previous study of about 1000 words (Bradley & Lang, 1999) and served as a validity check. A participant whose ratings of these words did not correlate sufficiently with those in the Bradley and Lang (1999) corpus was dropped from the analysis.

Prior to beginning the procedure, raters of pleasantness read the following instructions:

You are invited to take part in the study that is investigating emotion, and concerns how people respond to different types of words. You will use a scale to rate how you felt while reading each word....The scale ranges from 1 (happy) to 9 (unhappy). At one extreme of this scale, you are happy, pleased, satisfied, contented, hopeful. When you feel completely happy you should indicate this by choosing rating 1. The other end of the scale is when you feel completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored. You can indicate feeling completely unhappy by selecting 9. The numbers also allow you to describe intermediate feelings of pleasure... If you feel completely neutral, neither happy nor sad, select the middle of the scale (rating 5). Please work at a rapid pace and don't spend too much time thinking about each word. Rather, make your ratings based on your first and immediate reaction as you read each word. (Warriner et al., 2013, p. 1193).

For raters of arousal, the scale was described as ranging from 1 = "excited" [elaborated as "stimulated, excited, frenzied, jittery, wide-awake, or aroused"] to 9 = "calm" [elaborated as, "relaxed, calm, sluggish, dull, sleepy, or unaroused"] (p. 1193).

Consistent with practices in the psycholinguistic literature, Warriner et al. (2013) reported their results with ratings in reverse-scored format, such that 1 = unhappy/unpleasant [calm/unaroused] and 9 = happy/pleasant [excited/aroused]. We did the same.

## References

- Backer, T. E., Liberman, R. P., & Kuehnel, T. G. (1986). Dissemination and adoption of innovative psychosocial interventions. *Consulting and Clinical Psychology, 54*, 111–118. doi:10.1037/0022-006x.54.1.111.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*, 91–97. doi:10.1901/jaba.1968.1-91.
- Bailey, J. S. (1991). Marketing behavior analysis requires different talk. *Journal of Applied Behavior Analysis, 24*, 445–448. doi:10.1901/jaba.1991.24-445.
- Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs: General and Applied, 76*, 1–36. doi:10.1037/h0093918.
- Becirevic, A., Critchfield, T. S., & Reed, D. D. (2016). On the social acceptability of behavior-analytic terms: crowdsourced comparisons of lay and technical language. *The Behavior Analyst. Advance online publication*. doi:10.1007/s40614-016-0067-4.
- Berger, M. (1973). Behaviorism in twenty-five words. *Social Work, 18*, 106–108.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): stimuli, instructions manual and affective ratings (technical report no. C-1)*. Gainesville, FL: University of Florida, NIMH Center for Research in Psychophysiology.
- Buckley, K. W. (1982). The selling of a psychologist: John Broadus Watson and the application of behavioral techniques to advertising. *Journal of the History of the Behavioral Sciences, 18*, 207–221. doi:10.1002/1520-6696(198207)18:3<207::aid-jhbs2300180302>3.0.co;2-8.
- Coleman, C. (2011). Teaching health care professionals about health literacy: a review of the literature. *Nursing Outlook, 59*, 70–78. doi:10.1016/j.outlook.2010.12.004.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.
- Corrigan, J. D., Dell, D. M., Lewis, K. N., & Schmidt, L. D. (1980). Counseling as a social influence process: a review. *Journal of Counseling Psychology, 27*, 395–441. doi:10.1037/0022-0167.27.4.395.
- Dodds, P. S., & Danforth, C. M. (2009). Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *Journal of Happiness Studies, 11*, 441–456.
- Dodds, P. S., et al. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Science, 112*, 2389–2394. doi:10.1073/pnas.1411678112.
- Doughty, A. H., Holloway, C., Shields, M. C., & Kennedy, L. E. (2012). Marketing behavior analysis requires (really) different talk: a critique of Kohn (2005) and a(n)other call to arms. *Behavior and Social Issues, 21*, 115–134. doi:10.5210/bsi.v21i0.3914.
- Floh, A., Koller, M., & Zauner, A. (2013). Taking a deeper look at online reviews: asymmetric effect of valence intensity on shopping behavior. *Journal of Marketing Management, 29*, 646–670. doi:10.1080/0267257x.2013.776620.
- Fox, R. M. (1990). Suggested common North American translations of expressions in the field of operant conditioning. *The Behavior Analyst, 13*, 95–96.
- Fox, R. M. (1996). Translating the covenant: the behavior analyst as ambassador and translator. *The Behavior Analyst, 19*, 147–161.
- Freedman, D. H. (2015). Improving the public perception of behavior analysis. *The Behavior Analyst, 1–7*. doi:10.1007/s40614-015-0045-2.
- Garcia-Marques, T., Mackie, D. M., Claypool, H. M., & Garcia-Marques, L. (2004). Positivity can cue familiarity. *Personality and Social Psychology Bulletin, 30*, 585–593. doi:10.1177/0146167203262856.
- Guerini, M., Pepe, A., & Lepri, B. (2012). Do linguistic style and readability of scientific abstracts affect their virality? Download from <http://arxiv.org/pdf/1203.4238.pdf>.
- Howarth, P. (1994). *Attila, King of the Huns: the man and the myth*. New York: Barnes & Noble Books.
- Jarmolowicz, J. P., Kahng, A., Invarsson, E. T., Goysovich, R., Heggemeyer, R., & Gregory, M. K. (2008). Effects of conversational versus technical language on treatment preference and integrity.

- Intellectual and Developmental Disabilities*, 46, 190–199. doi:10.1352/2008.46:190-199.
- Lindsley, O. R. (1991). From technical jargon to plain English for application. *Journal of Applied Behavior Analysis*, 24, 449–458. doi:10.1901/jaba.1991.24-449.
- Maurice, C. (1993). *Let me hear your voice: a family's triumph over autism*. New York: Random House.
- Norman, S. M., Avolio, B. J., & Luthans, F. (2010). The impact of positivity and transparency on trust in leaders and their perceived effectiveness. *The Leadership Quarterly*, 21, 350–364. doi:10.1016/j.leaqua.2010.03.002.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, 23, 184–188. doi:10.1177/0963721414531598.
- Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic sound activation in silent word reading: evidence from backward masking. *Journal of Memory and Language*, 27, 59–70. doi:10.1016/0749-596x(88)90048-4.
- Petty, R. E., Schumann, D. W., Richman, S. S., & Strathman, A. J. (1993). Positive mood and persuasion: different roles for affect under high- and low-elaboration conditions. *Journal of Personality and Social Psychology*, 64, 5–20. doi:10.1037/0022-3514.64.1.5.
- Reagan, A., Tivnan, B.M. Williams, J.R., Danforth, C.M., & Dobbs, P.S. (2015). Benchmarking sentiment analysis methods for large-scale texts: a case for using continuum-scored words and word-shift graphs. Download from <https://arxiv.org/abs/1512.00531>.
- Rogers, E. M. (2003). *Diffusion of innovations*. New York: Simon & Schuster.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6, 412–415. doi:10.1111/j.1939-0025.1936.tb05248.x.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Press Press.
- Skinner, B. F. (1988). Reply to Harnad. In A. C. Catania & S. Harnad (Eds.), *The selection of behavior: the operant behaviorism of B.F. Skinner: comments and consequences* (pp. 468–473). Cambridge: Cambridge University Press.
- Smith, J. M. (2015). Strategies to position behavior analysis as the contemporary science of what works in behavior change. *The Behavior Analyst*, 1–13. doi:10.1007/s40614-015-0044-3.
- Warriner, A. B., & Kuperman, V. (2015). Affective biases in English are bi-dimensional. *Cognition and Emotion*, 29, 1147–1167. doi:10.1080/02699931.2014.968098.
- Warriner, A. B., Kuperman, V., & Brysbeart, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, 45, 1191–1207. doi:10.3758/s13428-012-0314-x.
- Witt, J. C., Moe, G., Gutkin, T. B., & Andrews, L. (1984). The effect of saying the same thing in different ways: the problem of language and jargon in school-based consultation. *Journal of School Psychology*, 22, 361–367. doi:10.1016/0022-4405(84)90023-2.
- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis*, 11, 203–214. doi:10.1901/jaba.1978.11-203.