In press: *Emotion*

Emotions *Do* Reliably Co-Occur With Predicted Facial Signals: Comment on Durán and Fernandez-Dols (2021)

Zachary Witkower[1]

Nicholas O. Rule[1]

Jessica L. Tracy[2]

[1]University of Toronto

[2]University of British Columbia

Please Address Correspondence to:

Zachary Witkower
Department of Psychology
University of Toronto
100 St. George Street
Toronto, ON Canada M5S 3G3
Zakwitkower@gmail.com

Abstract

Durán and Fernandez-Dols (2021) have done the field a service by conducting a meta-analytic review of the association between emotion experiences and facial expressions. Although they conclude that no meaningful association exists, our reading of their analyses suggest a different interpretation: The data that they report indicate an association of substantial magnitude—as large as one-and-a-half times the size of the average effect in social psychology, and larger than 76% of meta-analytic effects previously reported throughout personality and social psychology (Gignac & Szodorai, 2016; Richard et al., 2003). Moreover, reexamination of some of the exclusion and classification choices made by Durán and Fernandez-Dols (e.g., excluding intraindividual designs and studies purported to measure "amusement" from the primary analyses of "happiness") suggests that the observed large effects would be larger still if a more comprehensive set of studies had been included in their review. In sum, we conclude that Durán and Fernandez-Dols's meta-analyses provide robust evidence that emotions *do* reliably co-occur with their predicted facial signals, although this conclusion is opposite to the one stated in their report.

Emotions *Do* Reliably Co-Occur with Predicted Facial Signals: Comment on Durán and Fernandez-Dols (2021)

A defining feature of emotion is the coordination of experience and behavior (Ekman, 1972, 1992; Lazarus, 1991; Levenson, 1994; Scherer, 1984; Tomkins, 1962; Mauss et al. 2005). The field's emphasis on a coordinated system has been a springboard for more than 50 years of rigorous research testing for co-occurrence between subjective emotional experiences and facial expressions. Synthesizing this literature, Durán and Fernandez-Dols (2021) conducted five meta-analyses of existing data addressing the question of whether facial expressions co-occur with emotion-eliciting contexts and experiences. Their first three analyses, comprising the primary focus of the paper, examined basic emotions of anger, disgust, fear, happiness, sadness, and surprise. Analysis 1 included only studies on "whole" expressions (which resulted in the exclusion of all but happiness, surprise, and disgust), Analysis 2 included studies on "partial" expressions that examine core expressive features, and Analysis 3 included studies on partial expressions while also considering variation in emotion experience and intensity. In Analyses 1 and 2, Durán and Fernandez-Dols reported correspondences between facial expression and emotion contexts that ranged from p_{rp} = .13, 95% CI [.06, .27] to p_{rp} = .23, 95% CI [.13, .36].[1] In Analysis 3, the overall correspondence between facial expression intensity and the intensity of emotion experiences was r = .30, 95% CI [.18, .41].

The two remaining meta-analyses reported in the paper (Analyses 4 and 5), which the authors consider secondary analyses, focused exclusively on "amusement." Analysis 4 included studies that examined "partial" expressions of amusement (typically through experimental manipulations of the stimuli or context) and Analysis 5 included "partial" expressions of amusement while also considering variation in emotion experiences and intensity. Here, Durán and Fernandez-Dols reported

[1] Here, p_{rp} can be roughly understood as the proportion of participants assumed to be experiencing an emotion who also demonstrate the expression predicted for that emotion.

conventionally large correspondences of p_{rp} = .60, 95% CI [.38, .79] (Analysis 4) and r = .41, 95% CI [.34, .49] (Analysis 5). Based on the totality of their results, the authors conclude that "basic emotions do not reliably co-occur with their predicted facial signal" (p. 2). This conclusion is at odds with the data presented.

Although all of the reported effects indicate some degree of co-occurrence, the effect size reported in Analysis 3 is particularly informative for two reasons. First, both emotion experience and emotion expression are best conceptualized and measured as continuous, and variation in emotion intensity is expected to index facial expression intensity (Durán and Fernandez-Dols, 2021; Ekman, 1993). Falsely dichotomizing continuous variables can truncate effect-size estimates (e.g., Ferguson, 2009). Second, Analysis 3 provides an effect size metric (i.e., Pearson correlation) that can easily be compared to effects previously reported in the literature. The correlation between emotion experience and expected expression reported in Analysis 3 is r = .30, a medium effect by Cohen's (1988) standards and substantially larger (nearly 150%) than the average effect size across personality and social psychology (r = .21, Richard et al., 2003; also see Fraley & Marks, 2007; Gignac & Szodorai, 2016).

To be more precise, the correlation that Durán and Fernandez-Dols (2021) report in Analysis 3 is larger than roughly 76% of all published meta-analytic effects in social psychology (Gignac & Szodorai, 2016; Richard et al., 2003). For comparison, this association is similar in size to the association between self-reported extraversion and the proportion of time spent talking (r = .30, Tackman et al., 2021). It is roughly 1.5 times larger than the effects of meditation on anxiety, depression, and pain (rs < .19; Goyal et al. 2014), the association between employment interviews and job success (r = .20; McDaniel et al., 1994), and the effect of a nicotine patch (vs. placebo) on smoking abstinence (r = .18; Fiore et al., 1994). It is roughly twice as large as the associations between nonsteroidal anti-inflammatory drugs (e.g., ibuprofen) and pain (r = .14; Meyer et al. 2001), between

school grades and job performance (r = .16; Roth et al., 1996), between relapse prevention and improvement in substance abusers (r = .14; Irvin et al., 1999), between combat exposure and developing Post-Traumatic Stress Disorder (PTSD) within 18 years (r = .11; Centers for Disease Control Vietnam Experience Study, 1988), and between aggression and both testosterone levels and playing violent video games ($rs \leq .14$; Book et al., 2001; Sherry, 2001; Wood et al., 1991). Given this context, a more accurate conclusion from Durán and Fernandez-Dols' data would be that the co-occurrence between emotion experience and facial expression is similar in size or larger than other widely accepted and robust effects within the topic's shared field of social psychology.

Nonetheless, not all of the effects obtained by Durán and Fernandez-Dols (2021) are quite so large. For example, the proportion of participants placed in happiness, surprise, and disgust contexts who demonstrated the "whole" facial expressions predicted for those emotions is modest (p_{rp} = .13, 95% CI [.06, .27]; see their Analysis 1). However, this analysis on "whole" facial expressions is highly conservative and uncharacteristic of how facial expressions are usually conceptualized in emotion communication research.

Specifically, Duran and Fernandez-Dols define a "whole" facial expression as "a facial display that includes all the facial movements predicted by Ekman and Friesen (e.g., Ekman & Friesen, 1978; Ekman et al., 2002)" (p. 1551), referring to the Facial Action Coding System (FACS). Yet, in the FACS, Ekman et al. (2002) offer multiple variations of prototypes and multiple partial expressions ("major variants") for nearly every emotion (see Figure 10-1 from the FACS Investigators Guide for a detailed list; also see Kendler et al., 2008). Furthermore, Ekman (1992b) directly stated, "Often the combination of more than one muscle movement is necessary to clearly signal a single emotion, but that is not always the case" (p. 551). Ekman thus explicitly diffuses the notion that only "whole" or invariant facial expressions should be expected to co-occur with corresponding emotion experiences,

instead recognizing that emotion expressions—like almost all human phenomena—vary in meaningful ways, and that partial expressions are likely to occur.

Further supporting the ubiquity of partial expressions, in EMFACS (Ekman et al., 1994; also see Friesen & Ekman, 1983)—a condensed version of FACS designed specifically for coding emotions from their facial expressions—Ekman and colleagues tasked coders to score specific action units or critical combinations of no more than three action units that can be expected to co-occur. Ekman et al. (1994) referred to these as "core AUs" and "core combinations" in reference to their pivotal relation to emotion expression, even though these combinations and single AUs are not considered to constitute "whole expressions." In fact, we were unable to find any publication in which Ekman or his colleagues suggest that an expression must be "whole" to be scored as conveying a particular emotion; instead, by explicitly defining partial expressions in FACS, authoring EMFACS, and recognizing partial expressions in other written work, Ekman has proactively highlighted the existence and value of partial expressions.

Analysis 1 is therefore not only notably stringent and conservative but, by treating the extent to which individuals demonstrate perfectly prototypical, comprehensive, and invariant facial expressions in response to an emotion elicitor as the best test of co-occurrence, Analysis 1 is also a strawman test of the basic-emotions account. Indeed, the inclusion criteria for Analysis 1 were so stringent that studies examining only three of the six basic emotions could be included, thereby excluding all studies in more than 50 years of active research on anger, fear, and sadness. As a result of these criteria, Duran and Fernandez-Dols's (2021) Analysis 1, and the conclusions they draw from it, applies only to a subset of three emotions. We therefore disagree with Durán and Fernandez-Dols's assertion that Analysis 1 is "most theoretically relevant" (p. 6) and instead view Analyses 2 and 3 (which focus on the co-occurrence between emotion experiences and core expressive features) as more appropriate tests of whether emotion experience co-occurs with expected facial expressions.

Moreover, other decisions made by Duran and Fernandez-Dols (2021) likely underestimated the true effect size calculated from their primary analyses (i.e., Analyses 1-3). The association between happiness experiences and Duchenne smiles in Analyses 1-3 (and the resulting overall association between emotion experience and facial expression, broadly construed) excluded several studies with large effect sizes because, according to Duran and Fernandez-Dols, these studies measured "amusement" rather than "happiness." Although happiness rests at the conceptual core of amusement (Weidman & Tracy, 2020), and the original authors of the studies meta-analyzed did not indicate that the positive emotion they examined was amusement, Durán and Fernandez-Dols classified them as amusement, and relegated them to their secondary analyses (Analyses 4-5) on that basis, because the studies involved "humor elicitors" (p. 1561). However, our reading of the original studies and the emotion-elicitation contexts that their authors examined suggests that this classification was not always appropriate. Most notably, several of the meta-analyzed studies examined responses to success in a competitive situation; for example, Mui et al. (2017) measured children's facial behavior in response to winning or losing a competition, and uncovered a large-sized co-occurrence. Durán and Fernandez-Dols determined that the emotion measured in this study was amusement, although the word "amusement" appears nowhere in the article and the study contains no humor elicitor.

Similarly, two other studies relegated to secondary analyses for containing "humor elicitors" examined facial behavior produced by judo athletes after winning a competition, a situation that is not typically considered to provoke humor (nor did these studies' authors suggest that the emotion experienced by their participants was amusement; see Matsumoto & Willingham, 2006, 2009). Yet, two other studies included in Duran and Fernandez-Dols' meta-analyses measured facial behavior occurring in the *identical context* of winning a judo match (Crivelli et al., 2015) yet were classified by Duran and Fernandez-Dols (2021) as *not* measuring amusement, and thus included in their primary analyses (Analyses 1-3). No reason is given for this discrepancy, but it is clearly problematic to deem

two sets of studies measuring the same nonverbal displays shown in response to the same context as measuring two different emotions. Moreover, this decision was not trivial: Crivelli and colleagues' studies (which Fernandez-Dols co-authored) produced much smaller effect sizes than the studies by Matsumoto and Willingham (2006, 2009), and if Matsumoto and Willingham's studies had been included in the primary analyses, they would have constituted the largest effect sizes for happiness observed in Analyses 1-2, thus yielding a much larger overall effect. Needless to say, relegating studies that produced some of the largest effects to a secondary analysis because they either purportedly test a different emotion but actually do not (or because they allegedly result from a humor elicitation despite no indication of humor in the methods) is problematic and consequential for determining the co-occurrence between happiness and smiles, as well as between emotion experience and expressions more broadly.

This also raises a broader question: Why were amusement and happiness analyzed separately? If happiness expressions are defined based on Ekman's work, as Durán and Fernandez-Dols (2021) state, it would seem sensible to use Ekman's definition of happiness to classify studies in the meta-analysis. As Durán and Fernandez-Dols note, Ekman has repeatedly stated (and found) that amusement experiences should elicit happiness expressions (also called the Duchenne Smile; Ekman et al. 1990; also see Ekman et al., 1980; Ekman & Friesen, 1982; Ekman, 1992b). For example, Ekman et al. (1992) explicitly stated: "We consider happiness and enjoyment as general terms that cover a variety of different positive emotional experiences such as sensory pleasure, relief, satisfaction with accomplishment, amusement, and contentment. Although we expect that each of these positive emotions is experienced differently, with different sensations and physiology, we (Ekman & Friesen, 1982) have hypothesized that all of these positive states share the same expressive signal—the Duchenne smile" (p. 347). Ekman and colleagues therefore clearly regard happiness and amusement as feeling states conveyed through a shared facial expression.

Amusement and happiness overlap conceptually as well: A recent study found that subjective experiences of amusement, happiness, joy, elation, and contentment are empirically redundant with one another (Weidman & Tracy, 2020). Given the high degree of shared content and variance in both the experience and expression of amusement and happiness, it would make sense to include studies manipulating amusement in any analyses examining the co-occurrence between happiness and Duchenne smiles. Indeed, the finding that the co-occurrence between amusement and Duchenne smiles is remarkably high in Durán and Fernandez-Dols's (2021) Analyses 4 and 5 is itself evidence of the experiential and expressive overlap between amusement and happiness, as well as of the co-occurrence between happiness and its predicted expression.

Finally, all five analyses included correlations between emotion experience and predicted expression for effects *across* participants only (i.e., interindividual correlations). None of the five analyses included the potentially more theoretically relevant effects of correlations between emotion experience and expression *within* participants (i.e., from intraindividual designs); these were systematically excluded. A previous meta-analysis on the same topic, written by the same authors (Durán et al., 2017), included studies using intraindividual designs, and referred to them as "theoretically more adequate" (p. 3) and as yielding "higher coherence estimates than interindividual designs" (p. 17). Indeed, intraindividual designs constituted some of the largest effect sizes characterizing the association between emotion experience and expected expressions in prior research (e.g., Brown & Schwartz, 1980; Deckers et al., 1987; Durán et al., 2017; Fiacconi & Owen, 2015; Mauss et al. 2005; Mauss et al. 2011; Ruch, 1995). Given that individuals can have different emotion experiences (or different changes in emotion experience from baseline) in response to different elicitors (thereby potentially yielding different expressions), considering within-person variation provides a cleaner test of the co-occurrence between emotion experience and expression than between-person variation does. Surprisingly, these prior findings are not mentioned in the current paper.

Although we disagree with Durán and Fernandez-Dols's (2021) assertion that "basic emotions do not reliably co-occur with their predicted facial signal" (p. 2) and instead view the observed association as medium-to-strong in size, we also recognize that emotion experiences and emotion expressions do not match exactly. It is therefore important to note, in no uncertain terms, that variability in emotion expression exists. In the same way that two people who contract the same variant of the COVID-19 virus on the same day from the same person might exhibit unique symptoms (one with a runny nose but no sore throat, and one with a sore throat but no runny nose), two people experiencing the same emotion (e.g., surprise) might exhibit idiosyncratic facial behaviors (one with the eyebrows raised but no opening of the mouth, and one with opening of the mouth but no perceptible raising of the eyebrows). Variability in which coronavirus symptoms manifest for specific people does not mean that there is no reliable family of COVID-19 symptoms; likewise, variability in emotion expression does not mean that emotions have no reliable cluster of behaviors that convey a particular feeling state. Instead, the studies meta-analyzed by Durán and Fernandez-Dols indicate that one's subjective emotion experience is as likely to co-occur with a predicted facial expression as are many other medium-to-strongly related social psychological phenomena.

There are numerous reasons for variability in the co-occurrence between emotion experience and emotion expression, including (but not limited to) (a) similar experimental contexts eliciting somewhat different emotion experiences across participants; (b) cultural accents contributing to nuances in expression (e.g., Marsh et al., 2003; Elfenbein, 2003); (c) certain expressions being more readily expressed via bodily behaviors than facial behavior (e.g., Witkower & Tracy, 2018; Witkower, Hill et al., 2021; Witkower, Tracy et al., 2021); and (d) differences in emotion regulation or expression resulting from societal pressures, lived experiences, or other individual differences. In short, not all people behave identically in the same context. Instead, as is the case for perhaps all human behavior,

variability exists. Yet, the existing empirical evidence strongly suggests that emotions reliably co-occur with their predicted facial signal.

References

Book, A. S., Starzyk, K. B., & Quinsey, V. L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, *6*(6), 579-599.

Brown, S. L., & Schwartz, G. E. (1980). Relationships between facial electromyography and subjective experience during affective imagery. *Biological psychology*, *11*(1), 49-62.

Centers for Disease Control. (1988). Vietnam experience study. Health status of Vietnam veterans: I. Psychosocial characteristics. *JAMA*, *259*, 2701-2707.

Cohen, J. (1988). Statistical power analysis Jbr the behavioral. *Sciences. Hillsdale (NJ): Lawrence Erlbaum Associates*, 18-74.

Crivelli, C., Carrera, P., & Fernández-Dols, J. M. (2015). Are smiles a sign of happiness? Spontaneous expressions of judo winners. *Evolution and Human Behavior*, *36*(1), 52-58.

Deckers, L., Kuhlhorst, L., & Freeland, L. (1987). The effects of spontaneous and voluntary facial reactions on surprise and humor. *Motivation and Emotion*, *11*(4), 403-412.

Durán, J. I., & Fernández-Dols, J. M. (2021). Do emotions result in their predicted facial expressions? A meta-analysis of studies on the co-occurrence of expression and emotion. *Emotion (Washington, DC)*, *21*(7), 1550-1569.

Durán, J. I., Reisenzein, R., & Fernández-Dols, J. M. (2017). Coherence between emotions and facial expressions. *The science of facial expression*, 107-129.

Ekman, P. (1978). Facial Action Coding System. *Consulting Psychologists Palo Alto*.

Ekman, P. (1992a). An argument for basic emotions. *Cognition & emotion*, *6*(3-4), 169-200.

Ekman, P. (1992b). Are there basic emotions?. *Psychological Review*, *99*(3), 550-553.

Ekman, P. (1993). Facial expression and emotion. American Psychologist, 48, 384–392.

Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile: emotional expression and brain physiology. II. *Journal of Personality and Social Psychology*, *58*(2), 342-353.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of nonverbal behavior*, *6*(4), 238-252.

Ekman, P., Friesen, W. V., & Ancoli, S. (1980). Facial Signs of Emotional Experience. *Journal of Personality and Social Psychology*, *39*(6), 1125-1134.

Ekman, P., & Friesen, W. V. (1982). Felt, false, and miserable smiles. *Journal of Nonverbal Behavior*, *6*(4), 238-252.

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial Action Coding System: Facial action coding system: the manual*. Research Nexus.

Ekman, P., Irwin, W., & Rosenberg, E. (1994). EMFACS-8: Coders Instructions (EMFACS-8). Univ. California San Francisco Press, San Francisco, CA.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532-538.

Fiacconi, C. M., & Owen, A. M. (2015). Using psychophysiological measures to examine the temporal profile of verbal humor elicitation. *PLoS One*, *10*(9), e0135902.

Fraley, R. C., & Marks, M. J. (2007). The null hypothesis significance testing debate and its implications for personality research. *Handbook of research methods in personality psychology*, 149-169.

Friesen, W. V., & Ekman, P. (1984). EMFACS-7: Emotional facial action coding system, Version 7. *Unpublished manual*.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and individual differences*, *102*, 74-78.

Goyal, M., Singh, S., Sibinga, E. M., Gould, N. F., Rowland-Seymour, A., Sharma, R., ... & Haythornthwaite, J. A. (2014). Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. *JAMA internal medicine*, *174*(3), 357-368.

Kendler, K. S., Halberstadt, L. J., Butera, F., Myers, J., Bouchard, T., & Ekman, P. (2008). The similiarity of facial expressions in response to emotion-inducing films in reared-apart twins. *Psychological medicine*, *38*(10), 1475-1483.

Matsumoto, D., & Willingham, B. (2006). The Thrill of Victory and the Agony of Defeat: Spontaneous Expressions of Medal Winners of the 2004 Athens Olympic Games. *Journal of Personality and Social Psychology*, *91*(3), 568-581.

Matsumoto, D., & Willingham, B. (2009). Spontaneous Facial Expressions of Emotion of Congenitally and Noncongenitally Blind Individuals. *Journal of Personality and Social Psychology*, *96*(1), 1-10.

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The Tie That Binds? Coherence Among Emotion Experience, Behavior, and Physiology. *Emotion*, *5*(2), 175-190.

Mauss, I. B., Shallcross, A. J., Troy, A. S., John, O. P., Ferrer, E., Wilhelm, F. H., & Gross, J. J. (2011). Don't hide your happiness! Positive emotion dissociation, social connectedness, and psychological functioning. *Journal of personality and social psychology*, *100*(4), 738-748.

McDaniel, M. A., Whetzel, D. L., & Schmidt, F. L. (1994). The validity of employment interviews: a comprehensive review and meta-analysis. *Journal of Applied Psychology*, *79*(4), 599-616.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., ... & Reed, G. M. (2001). Psychological testing and psychological assessment. A review of evidence and issues. *The American psychologist*, *56*(2), 128-165.

Mui, P. H. C., Goudbeek, M. B., Swerts, M. G. J., & Hovasapian, A. (2017). Children's nonverbal displays of winning and losing: Effects of social and cultural contexts on smiles. Journal of Nonverbal Behavior,41(1), 67–82.

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of general psychology*, *7*(4), 331-363.

Ruch, W. (1995). Will the real relationship between facial expression and affective experience please stand up: The case of exhilaration. *Cognition & Emotion*, *9*(1), 33-58.

Sherry, J. L. (2001). The effects of violent video games on aggression: A meta-analysis. *Human communication research*, *27*(3), 409-431.

Weidman, A. C., & Tracy, J. L. (2019). Picking up good vibrations: Uncovering the content of distinct positive emotion subjective experience. *Emotion*, *20*(8), 1311.

Wood, W., Wong, F. Y., & Chachere, J. G. (1991). Effects of media violence on viewers' aggression in unconstrained social interaction. *Psychological bulletin*, *109*(3), 371-383.