Chapter 3.2

On the use of actor portrayals in research on emotional expression

Scherer, K.R. & Bänziger, T.

Note: titles and text are subjected to change, do not quote verbatim

Abstract

In this chapter we address the hotly debated issue of the utility of using actor portrayals in research on emotional expression in psychology, neuroscience, and affective computing. We argue that emotions are rare and fleeting events that are difficult to capture in a purely spontaneous fashion, especially as they are likely to be constantly manipulated for the purpose of self-regulation or social constraints. We propose an account of the fundamental mechanism underlying emotional expression and present a theoretical analysis that distinguishes between push (physiologically driven) and pull (social regulation and strategic intention) factors, suggesting that corpora with unobtrusive recordings of real-life expressions, laboratory induction, recording of expressive behavior from media shows, and explicit actor portrayals all have their place in EE research and vary continuously on several dimensions rather than representing completely different classes of expressions. We argue that an analysis of pure push factors, spontaneous unregulated expressions, is unrealistic in practice and probably of little interest, given the scarcity of such pure expressions in social life. We suggest that one focus of current research should instead be directed toward the explicit study of pull effects, the use of actor portrayals being a highly appropriate methodological choice for this aim, given the possibility of manipulating and standardizing pull effects. The central role of actor portrayals clearly lies in the empirical and experimental study of the shared code of emotional signaling and the examination of cue utilization in emotion perception and inference.

The study of emotional expression (EE) plays a central role in emotion research as visible or audible expressions externalize an internal state and thus become a major aspect of social communication, informing others about the person's reactions to events and intentions to act. Not surprisingly, then, EE research constitutes a sizeable proportion of emotion research in psychology, ethology, neuroscience, and affective computing. A major problem for such research efforts is the fact that strong emotions that are likely to be forcefully expressed are relatively rare, fleeting phenomena that are generally hidden from public view or covered up. The fact that strong, prototypical emotions are relatively rare is suggested by indirect evidence. Thus, when one tries to report a representative instance of certain emotion experiences, for example intense sadness, one often has to go back in time for several months to find an appropriate example in memory (Scherer & Wallbott, 1994, p. 319). Only relatively frequent emotional experiences, such as minor anger episodes, come readily to mind by recalling events from previous days or weeks. In an actuarial study of emotional experiences, asking a quasi-representative sample of the Swiss population to report the most important emotion that they experienced yesterday, only anger or joy were reported by more than 10% of the sample whereas emotions such as sadness or fear were reported only by about 2-3%, suggesting that such strong emotions occur rather less frequently in daily life (see Scherer, K., Wranik, Sangsue, Tran, & Scherer, U., 2004).

Even EEs that occur relatively frequently in everyday life are often quite unpredictable, which makes them hard to catch for the researcher. EEs are fleeting, evanescent events that come and go rapidly. This means that if one wants to go beyond impressionistic description and interpretation, EEs need to be recorded audio-visually at sufficiently high quality to allow objective measurement and analysis.

Unfortunately, apart from lack of opportunity, that is, the difficulty of catching appropriate incidents of EEs in a naturalistic fashion in the field, EEs are also almost

impossible to systematically observe or even record for further analysis because of privacy and ethics constraints. Even if it were possible to obtain the informed consent of all concerned, the recording conditions (camera angle, image resolution, sound quality) are often hampered by the available technology or the local setting. Even more important, generally only public settings provide ethically admissible observation and recording options, and so the likelihood is strong that senders will closely monitor and control their expressive behavior, conforming to sociocultural display rules (Ekman & Friesen, 1971; Matsumoto, 1990). This tendency will be even more pronounced in the presence of cameras and microphones, producing the danger that the authentic expression of the respective emotions will be completely suppressed or modified and sometimes replaced by stereotypic expression patterns that are in accordance with cultural expectations for the respective event and setting.

An alternative approach consists of inducing an emotion in the laboratory and observing or recording the corresponding expression. Although psychologists have created an impressive number of induction techniques (Coan & Allen, 2007), some of which produce relatively reliable effects, the intensity of the resulting states is generally low, with little outwardly observable expression, for example in the case of emotion induction through picture viewing, listening to music, or imagination and memory retrieval. Ethical constraints, as well as cost and practicality, often prevent researchers from confronting participants with stimuli or events of sufficiently high importance or relevance, that are likely to produce bona fide emotions. Most important, expressions are often generated by action tendencies (Darwin, 1872/1998; Frijda & Tscherkassof, 1997; Scherer & Ellgring, 2007), and because most experimental inductions and manipulations in the laboratory do not require or allow adaptive action, the potential for the production of EEs is limited. Furthermore, as in the case of observation in natural field settings, artifacts cannot be excluded. Thus, by using social or personal display rules, participants may suppress or modify naturally occurring expressions

that they consider inappropriate, or may even simulate emotions that they do not experience to show that they are "good participants" who give the experimenter what is expected.

In sum, neither natural field observation nor experimental induction are particularly conducive to EE research -- neither for the study of the mechanisms underlying EE production or of the way in which they are perceived and interpreted by observers. At the same time, the methodological demands on this research with respect to experimental design and procedures are extensive as the following issues have to be taken into account:

1. Individual differences -- The evidence for strong individual differences in emotional reactions is abundant. Different individuals (or the same individual at different times) may react with different emotions to the same situation because differentially salient goals and values that are specific to an individual may lead to different appraisals (Scherer & Brosch, 2009). Individuals also have different response or coping styles (including expression; e.g., externalization vs. internalization). For this reason, expressions of several different emotions and repeated expressions of each emotion need to be sampled for each individual studied. This sampling needs to occur under highly controlled situational conditions because even minor differences in event appraisal may lead to widely different emotions (Ellsworth & Scherer, 2003; Scherer, Schorr, & Johnstone, 2001).

2. Emotion differences: As mentioned earlier, many EEs are generated by action tendencies and because each emotion can generate a variety of partially overlapping action tendencies, the study of EE differences between emotions requires the study of widely different emotions. In consequence, studies on EE need to sample as large and as complete a set of emotions as possible.

3. Technical requirements: The need for objective coding and measurement of multimodal EEs requires audio-visual recording and a high level of sophistication to ensure a high recording quality required by the advanced annotation and analysis methods in this domain (see Harrigan, Rosenthal, & Scherer, 2005).

Clearly, these conditions are hard to meet, especially at reasonable cost in time and money, if one wants to record naturally occurring EEs for many different emotions, in a repeated fashion, for a sizeable number of comparable individuals in everyday life or to induce many different strong emotions in a repeated fashion in the laboratory. The limited suitability of field observation and experimental induction paradigms for the systematic study of EEs and the need for large-scale sampling, repetition, situational control, and sophisticated methodology have led expression researchers to take recourse in actor portrayals of emotion. It is no surprise that the bulk of expression research reported in the literature, particularly for the perception and interpretation of EEs, has used this approach, generating a remarkable body of pertinent and highly replicable findings.

The portrayal paradigm generally consists in asking trained lay persons or professional actors to produce an EE that can be considered a plausible or believable instance of a prototypical expression of a given emotion. In some cases, encoders are only given an emotion label and asked to produce the appropriate expression. More frequently, however, more elaborate encoding instructions are used, for example providing typical emotion scenarios which the actors are to act out, or a procedure that approaches induction procedures, namely the Stanislavski or method acting approach which requires the actors to generate a rudimentary feeling state corresponding to the respective emotion by using memory recall or vivid imagination techniques. Providing standard scenarios serves to define the emotion more precisely beyond the semantic meaning of words and is likely lead to more comparable portrayals across actors for a given procedure. In contrast, a Stanislavski procedure, through the implication of personal experiences and images, encourages somewhat more idiosyncratic expressions. Both procedures are likely to increase the authenticity and believability of the portrayals as they discourage the use of stereotypical patterns. In some cases, both approaches are combined for maximal effect (see chapter 6.1).

Unfortunately, actor portrayals are often misunderstood. The purpose of using actor portrayals is *not* to study spontaneously occurring emotions or to detect underlying emotions in actors. Actors are asked to produce expression patterns that are plausible and believable replicas of real, spontaneous expressions typical of certain emotions. Participants in judgment studies are asked to judge what emotion is *represented* by the portrayal, not what emotion the actor feels. And the purpose of the research is to identify the prototypical representation of emotions in social communication, not to study the nature of spontaneous emotion expressions in real life. Clearly there are differences between the two. For example, differences in timing and dynamic patterning have been shown between "true" and "fake" smiles despite the fact that the facial movement is the same (Krumhuber & Kappas, 2005; Schmidt, Ambadar, Cohn, & Reed, 2006). But the purpose of using actor portrayals in emotion perception research is to examine the expressive *code* used in social emotion communication.

Studying a code requires that it is understood by the decoder and thus portrayed expressions are generally selected on the basis of their recognizability, something which some critics fail to understand. Thus, Barrett (2006) writes: "For example, one very influential study (Banse & Scherer, 1996) attempted statistical classification of only 16.9% of the vocal samples that were actually recorded. It is difficult to interpret the findings from meta-analyses like the one by Juslin and Laukka (2003) when the utterances being classified are carefully selected to represent only a small subset of those that occur within the lab." There is a fundamental misunderstanding here. Studying a communication code requires the selection of well-recognized tokens; otherwise, the precondition of socially shared iconic representations is not met. Generally, in portrayal studies, many more repetitions are recorded than stimuli are needed and a selection of the most reliably rated and best recognized stimuli is made to ascertain that a shared iconic representation is captured.

In consequence, the utility of actor portrayals is intrinsically limited to the study of the shared code of EE in emotion communication and care has to be taken to avoid using such material in an inappropriate fashion, Clearly, EE corpora or databases need to be adapted to the specific research aims pursued by the researcher and a careful analysis of the specific needs should made in each case (see Chapter 3.1 this volume; Douglas-Cowie, Campbell, Cowie, & Roach, 2003). In many cases one may need to resort to different techniques to obtain pertinent samples of EE, using systematic induction, media material, field observation, or other techniques (see Chapter 6.2; Campbell, 2000; and

http://www.acii2009.nl/program/show_slot/42). However, care should be taken to avoid using a simple dichotomy between actor portrayals, considered as artificial and inauthentic, and corpora obtained through laboratory induction, from media records, or through surreptitious field recordings, considered as natural or authentic. We believe that it is necessary to examine the underlying conceptualization of the distinction between "natural and artificial" to evaluate this suggestion. The general assumption seems to be that expressions can be arranged with respect to a simple binary distinction:

- natural, realistic, authentic, genuine, real, believable, sincere, trustworthy

versus

- unnatural, unrealistic, counterfeit, artificial, fake, not believable, feigned, unreliable

We argue that this is an unrealistic claim as expressions tend to vary continuously on several dimensions. Thus, closer inspection of the above enumerations shows that these qualifiers take different objects and have different meanings. The following list provides a small sampling of different pairs of terms and suggests criteria that might be used for deciding which term of the pair should be applied:

Terms	Criteria
Natural vs. unnatural	Biological determination

Realistic vs. unrealistic	Probability of occurrence
Genuine vs. artificial	Original vs. copy
Real vs. fake	Original vs. copy
Authentic vs. counterfeit	Deception intention
Sincere vs. feigned	Deception intention
Believable vs. not believable	Observer judgment
Trustworthy unreliable	Observer judgment

Of course, this is a very hazardous procedure as many of the terms are semantically highly complex. Naturalness seems to refer to something that is biologically given and immutable, whereas something unnatural would go against preordained practice. The quality of being realistic, corresponding to facts rather than abstract ideals, could be glossed as the probability of occurrence of a specific behavior in a given situation determined by its actuarial frequency and prototypicality. Unrealistic would apply to figments of the imagination. Genuineness seems to depend on whether the object or behavior in question is the original or a copy, implying the existence of a certain uniqueness. Sincerity or authenticity can only be judged on the basis of a person's intention to deceive about a true underlying state, feigning or faking another. And believability or trustworthiness surely depends entirely on subjective observer judgment rather than on some objective quality of the behavior. From this analysis, we suggest distinguishing three underlying perspectives for classifying EEs:

1.A *behavior perspective* that examines the frequency and typicality of a given EE in the context of a specific situation, thus defining its realism or naturalness, as the biological determined should also be the most frequent. This perspective requires an examination of the biological determination of EE and actuarial investigation of frequency and typicality.

2.A *sender or encoder perspective* that examines the EE from the point of view of the production, taking into account communication intentions in particular. Here, issues such as

spontaneous versus voluntary, strategic production, raw versus regulated, controlled expression, and conscious versus unconscious production have a role. The true underlying intention is exceedingly difficult to determine, even by questioning the sender, because the production factors may be unconscious or the sender may want to dissimulate strategic intentions. Most important, both types of factors may be involved in close interaction.

3.A *decoder or observer* perspective that is orthogonal to the preceding factors because observers may form subjective impressions about prototypicality and thus realism, or about sender intention and thus authenticity, or about sincerity versus strategic intention. Genuine observer judgments are perceived believability or plausibility (an evaluation that combines realism and authenticity) and perceived reliability or trustworthiness of the sender (a derived judgment about sender dispositions).

Given these three perspectives, it is possible to analyze more precisely the differential utility of different types of sampling of EE as corpora or data bases for research, such as actor portrayal procedures, excerpts from media game or reality shows, laboratory induction, or surreptitious or open recording of spontaneous, real-life EE s. Which of these research paradigms come closest to the natural, authentic, and genuine expressions in spontaneous emotion expressions that are presumably biologically determined, untainted by devious encoder intentions, and thus automatically believable and trustworthy? Cowie, McKeown, & Gibney (2009 show, using examples in which spontaneous emotion is perceived as acted and acted emotions as spontaneous and authentic, that this question has no simple answer. They focus on the third perspective outlined above, concerned with observer judgment, and demonstrate the need to study the nature and distribution of expressive cues in multiple modalities that underlie authenticity judgments of observers. They note that the differences between the various types of sampling needed to be discussed with respect to their relative importance and the research questions and demands on analysis in specific cases. The perspectives we have outlined above can help to examine these issues. Thus, the first,

behavior perspective, is obviously central in cases in which there are clear context constraints and in which it thus makes sense to require high frequency and prototypicality of expressions. The third, decoder or observer perspective, is most important in research questions that examine the mechanisms underlying emotion perception, the nature of the cues, and the impression formation and inference processes (see also Figure 1). Both of these perspectives play a major role in the study of the interpersonal communication of emotion. The most difficult is the second, sender or encoder perspective, as it requires to identify the "rael state" and communication intentions of the sender which are often unconsciously or purposively hidden from public inspection. This perspective is of central relevance for diagnostic purposes in which the investigator wants to use EE to uncover the sender's innermost feelings and true intentions, for example in the case of detecting deception (Ekman, 2009). The problem is that the distinction between true and faked feeling is a very difficult one. We will devote the bulk of this chapter to address some of the underlying theoretical issues for this distinction.

A good starting point, as so often, is Aristotle, who, in remarkable anticipation of the notion of emotional competence or intelligence, pointed out that to qualify as a socially skilled person one needs "to be angry in the manner, at the things, and for the length of time, that the rule dictates" (see Aristotle, in McKeon, 1941, p. 996); in other words, anger expression has to conform to social standards and expectations. This theme was echoed and enlarged by Goffman's (1959) notion of impression management, in which one's emotions are expressed in accordance not only with sociocultural norms, but also with one's strategic intentions. Clearly, it is not enough to closely control EE; rather, the expression has to be appropriate for cultural and contextual expectations. As shown by work on *display rules* (Ekman, Sorenson, & Friesen, 1969; Matsumoto, 1990; Scherer, 2000), people need to actively produce the *appropriate* expression. These mechanisms are clearly incompatible with the pure natural, genuine, and authentic view espoused by many researchers.

So what exactly are the determinants of emotional expression? Scherer and his collaborators (Kappas, Hess, & Scherer, 1991; Scherer, 1985, 1986, 1988; Scherer, Helfrich, & Scherer, 1980; Scherer, & Kappas, 1988) have suggested to settle the conundrum by differentiating *push and pull effects* in EE. In the case of push effects, internal factors, such as physiological arousal in strong emotions, "push" motor behavior into certain directions (such as adaptive actions) but do not necessarily target particular configurations. Thus, the underlying emotion processes are highly variable and volatile; in consequence, the resulting expression is also highly variable and may rapidly change over time. In contrast, external pull factors are in the service of specific communication intentions or of culturally defined norms or expectations that require the production of relatively unambiguous expressive features in a specific signal structure. The sender needs to produce this pattern to achieve a particular effect. In this case, the expression outcome or target is fixed, or at least constrained, although the processes by which it is brought about can be variable.

Much research on EE has been conducted with an implicit focus on push factors in an attempt to use expression as direct readout of the underlying emotional state of the individual. This research is often frustrating because researchers rarely have access to bona fide emotions in natural social settings, individuals differ strongly in their expressive behavior, and it is difficult to objectively determine the precise nature of the underlying emotion without using verbal report (which is often biased by defense strategies or conventional response rules). In this chapter, we attempt to redress the balance and to focus on the pull effects in emotion expression and communication.

Mechanisms Underlying the Operation of Push and Pull Factors

Figure 1 shows an adaptation of the Brunswikian lens model that Scherer (1978, 2003) has suggested as a guide to theorizing and research in the area of EE. The model suggests that the expresser encodes an emotion by producing a number of distal cues in bodily posture, gestures, facial movements, and vocal cues. These distal cues are transmitted

via the auditory and visual communication channels to a perceiver, where they give rise to proximal percepts; the correspondence to the original distal cues depending on transmission quality and the capacity of the sense organs. In the process of impression formation, the perceiver uses inferential mechanisms to attribute an emotion to the expresser (with variable degrees of accuracy). Unfortunately, this model has been rarely used to describe EE and the mechanisms underlying both the generation of distal cues as a function of the underlying emotion on the one hand and the inference and attribution of emotion on the basis of the proximal percepts on the other (but see Bänziger, 2004; Juslin, 1998, 2001). In particular, most research has focused on single expression modalities such as facial or vocal cues, making it impossible to study the coherence between the different modalities that may have a major role in constituting the emotion (see Scherer, 2009) and in determining perceiver inference (see below).

_____

Figure 1 about here
_____

The proposed distinction between push and pull effects implies specific hypotheses about the nature of the underlying mechanisms. As shown in Figure 1, in the case of push factors, one expects the operation of neurobiological mechanisms that generate specific expression patterns as part of the emotion process. Push effects are biologically determined externalizations of naturally occurring internal processes of the organism, particularly information processing and behavioral preparation. Examples for cases in which push effects dominate are reactive animal expressions; infant grunts and cries; affect bursts; or sudden, uncontrolled emotional utterances.

What is the underlying mechanism? In his component process model, Scherer (1984, 1992, 2001, 2009) proposes that efferent effects of sequentially accrued appraisal results cumulatively constitute the unique, context- and individual-specific response pattern for a

given emotion episode. The component process model is based on the idea that during evolution, emotion has been optimized to serve the following functions: (a) evaluation of objects and events, (b) system regulation, (c) preparation and direction of action, (d) communication of reaction and behavioral intention, and (e) monitoring of internal state and organism–environment interaction (see Scherer, 1984, 2001). In consequence, the model predicts that the results of sequential appraisal checks will generate appropriate response patterns, including, particularly, the generation of the expression patterns. Predictions for facial, vocal, and gestural expressions have been elaborated (Johnstone, van Reekum, & Scherer, 2001; Kaiser & Wehrle, 2001; Scherer, 1986, 1987, 2001, 2009) on the basis of several classes of determinants: (a) the effects of the physiological change, (b) the preparation of specific instrumental motor actions, and (c) the production of socio-communicative signals. The first two determinants can be subsumed under the push effects.

As shown in Figure 1, push effect encoding is determined by neurobiological mechanisms. Of particular importance are three major instrumental functions of the facial organs (lips, nose, ears) and the vocal tract (mouth, pharynx, larynx): (a) passing matter (light, air, liquids, solids) to and from internal organs (e.g., in the service of respiration, metabolism, and glandular secretion); (b) positioning sensory organs for optimal reception of stimulation (e.g., raising eyebrows, flaring nostrils); and (c) acting directly on objects and other organisms (biting, licking, kissing). Table 1 in Scherer (2009) shows the predictions for facial, vocal, and gestural expressions resulting from individual appraisal checks.

The model suggests that the cumulative results of a sequential series of checks (1, relevance of the event; 2, implications for major needs, goals, and values; 3, ability to deal with these consequences or coping potential; and 4, normative significance of the event) produce a wide variety of complexly patterned emotion episodes. Despite this variability, a number of modal emotions, such as anger, fear, or joy, can be identified (Scherer, 1994b). Yet, one can expect relatively strong interindividual differences in the expressive patterns

produced by push effects (as the underlying biological processes are dependent on both the idiosyncratic nature of the individual and the specific nature of the situation).

In contrast, the social signaling function is served by pull effects, that is, particular visual or auditory signal configurations that are part of a socially shared communication code. This pull effect encoding is determined by linguistic rules for the encoding of syntactic, semantic, and pragmatic aspects of meaning and socioculturally variable norms, or molds, concerning the signal characteristics required by the shared codes for the communication of internal states and behavioral intentions. Examples for pure pull effects are sound symbolism, symbolic coding systems such as language, conventionalized expression rules, affect emblems, the mimicking of push effects, or constraints due to specific communication channels. Because pull effect encoding is characterized by a high degree of symbolization and conventionalization, one would expect comparatively few and small individual differences.

Just as push and pull effects differ in their underlying production or encoding mechanisms, they also differ in the corresponding perception or decoding mechanisms. As shown in Figure 1, one would expect animals and humans to have innate or prepared recognition schemata and efficient learning strategies for push effect expressions, whereas pull effects are likely to be decoded on the basis of socially transmitted, and explicitly taught, decoding rules that determine inference and attribution of the transmitted emotional meaning.

Pure push or pull effects are rare, because these two classes of determinants always closely interact (Kappas et al., 1991; Scherer, 1985, 1988; Scherer & Kappas, 1988). This is why Figure 1 illustrates the simultaneous effects of push and pull production and perception mechanisms in the ongoing process of emotion transmission. Once the emotion process has started, the resulting patterning becomes itself subject to appraisal (see also Frijda, 1993) and thus subject to different kinds of regulation. Furthermore, emotion, although originally elicited in a spontaneous fashion, may, in the ongoing process of emoting, become of strategic importance in interaction, as the claims by Aristotle (in McKeon, 1941) and

Goffman (1959), mentioned earlier, show. Thus, once the emotion process is ongoing, one expects various mixtures of biologically driven (often automatic) and socially learned (sometimes intentional) affect expression, depending on the nature of the situation, the strategic goals of the actor, and his or her cognitive and social development. In consequence, one of the major issues to be dealt with in the study of behavior control under emotion is the relative mix of push versus pull, or more spontaneous versus more deliberate, aspects of expression control.

One can assume that most real-life EEs have a strong pull effect component because much behavior in social settings is closely regulated and controlled. Thus, even in the case of the most primitive expressions, the affect bursts (Scherer, 1994a), one is much more likely to find them in a "domesticated" version (Wundt, 1900) in the form of affect emblems or interjections. Yet, many researchers in the domain of EE research pursue a quest to study the natural, true, authentic expression of the observed individual's innermost feelings, in other words, pure push effects. Apart from the problem that such cases are likely to be rare, they are unlikely to happen in the presence of research teams with the capacity for high-quality multimodal recordings of such instances. Much of the EE research has a focus on expressions that have a strong pull component, including signals that the individual wants his social environment to notice and interpret correctly. We believe that it is essential to systematically investigate the distal cues that are produced by senders in such cases and to examine how these cues are "utilized," in the Brunswikian sense, by decoders or receivers who try to make sense of the proximal cues they perceive. Given the difficulty of studying pure push effects in EE, and given the pervasiveness of cases with mixed origin, often with considerable pull effects, this investigation seems to be an important and realistic aim.

Furthermore, the existence of pull mechanisms implies that specific expressions are socially appropriate for specific senders in specific social situations and that observers (receivers of the displays) will interpret the expressions according to the context of their

production. Many examples can be used to illustrate this point, for instance, the reaction of the "winner" and the "loser" in the final of a beauty pageant. These contests almost always include a finale in which two candidates are singled out and where the tension and anticipation rises while everyone awaits the nomination of the winner. Rules for appropriate emotional displays and rules for the interpretation of the behavior of the contestants are well established. The loser is probably disappointed but does not display this emotion; instead showing happiness (for the sake of the winner) and good humor. The winner is mostly happy, but the display must also include tears, tremors, and signs of surprise or disbelief. Viewers are aware of this scenario; both contestants hope to win, yet both act as if they expected the other one to win (this rule seems to apply to other/competitions as well, but not to all types). Viewers would probably be shocked if the expected scenario was not fulfilled. However, someone foreign to the rules of this specific situation might not understand the scenario and might interpret the reactions in a different way. Likewise, in most social situations there are well-established rules that orient the interpretation of the nonverbal displays presented by specific social actors. Although the context is inseparable from the nonverbal displays in everyday experience, the ensuing confounding of context and nonverbal displays constitutes a problem for the advancement of research into EEs. When a nonverbal display (EE) is embedded in other contextual cues that provide information about the probable emotion(s) of the sender (e.g., in the example of beauty pageant, it would be impossible to witness the ongoing nonverbal reactions of the contestants without getting information about the underlying scenario at the same time), it becomes impossible to assess the specific contribution of the nonverbal displays to the formation of an impression in the viewers. To find out if and how the nonverbal displays contribute to the interpretation of the emotions experienced by the senders, one would need to isolate the nonverbal displays and to present them in different contexts or to present different nonverbal displays in the same context.

If we analyze media recordings of EEs from a push-pull perspective, one may conclude that many takes do not seem highly natural or realistic (except in their own contexts, e.g., TV game shows) and are unlikely to be devoid of strategic intentions (given the evaluation apprehension in the recording situation). Similarly, laboratory-induced EEs, if one could produce them in a reliable fashion, would suffer from some of the same shortcomings: the expressions might be specific to the experimental context and demand characteristics are likely to play a major role. Although media conditions may increase the chances that strategic impression intentions will have a role, the same is true in real life, as our review of the Goffmanian approach, highlighting the "presentation of emotion in everyday life," has shown (see also the writings of ancient rhetoric teachers such as Aristotle, Cicero, or Quintilian). Thus, the difference to explicit portrayals by actors may be much less important than is generally held. As argued above, all expressions of emotion vary continuously on the dimensions outlined above and we can expect gradients rather than categorical distinctions. Obviously, it is highly instructive to obtain corpora with media recordings, chance unobtrusive recordings from "real life", as well as to induce a variety of emotional states in the laboratory and observe the expressions. In every single case, there will be an interaction of push and pull effects and it would be of great benefit to pull these apart. What we want to warn against is to treat the difference to explicit actor portrayals as a fundamental categorical distinction that opposes the natural to the artificial.

Using Actor Portrayals in Systematic Research on Pull Factors

Rather, we argue that the use of actor portrayals of EEs provides a royal road to examine the questions outlined above in a principled fashion. In the last quarter of the 18th century, the French philosopher and encyclopedist Diderot wrote an essay on the "actor's paradox" ("Le paradoxe du comédien") in which he debated the question of whether and how actors can produce authentic, believable emotions on the stage without simultaneously

experiencing these emotions. The issue of the "player's passion" has been passionately debated ever since by both actors and scholars alike (see Roach, 1993) and it is of central relevance to contemporary emotion research. Recently, this debate has acquired new timeliness, given the growing dissension about the "right" way to study EE in voice, face, and body (Douglas-Cowie, Campbell, Cowie, & Roach, 2003). The use of actor portrayals has a long history in the study of emotion research because of the inherent limitations of studying naturally occurring EEs.

We have attempted to show that neither sampled media nor real-life EEs are devoid of strategic intentions leading to simulation, fabrication, control, and so forth. Furthermore, in many of these cases, the nature and strength of these intentions are unknown or extremely difficult to determine reliably. Thus, it may be overly optimistic to expect to obtain natural, authentic, spontaneous, sincere, and believable EEs through media sampling or laboratory induction. Furthermore, as long as one does not have an exact inventory of all the components of an emotional episode, including the person's conscious feeling state and label, one has no clue as to the authentic, natural emotion that is expressed; and it is dubious if such a state exists at all, because in all emotion episodes, control and regulation enter from the start.

How do actor portrayals fare in comparison? One central point is that the use of actor portrayals in research is not an attempt to study natural, authentic, spontaneous, sincere, and believable EEs. It is clear from the outset that the actor will intentionally produce an EE that may be completely unrelated to his or her current emotional state. The major advantage of actor portrayals, in contrast to media sampling or laboratory induction paradigms, is that the strategic intentions of the actor are known and can be standardized. The actor is told to produce a believable or plausible expression that will be recognized by observers and that must be prototypical and consequently natural and realistic. This, we hold, will make it possible to study the pull effects of EE and communication in the sense described earlier, because the actor's expressions are likely to come close to the templates that would be,

consciously or unconsciously, used in producing expressions that are "pulled" into the direction of a specific strategic or culturally imposed target.

However, we assert that to achieve satisfactory results, asking students or members of one's laboratory to "do" certain emotions is not sufficient. From Diderot on, the consensus has been that it is an art to produce believable emotions on the stage and that only the best professionals succeed in this enterprise. On the basis of theory and our own experience (Banse & Scherer, 1996; Gosselin, Kirouac, & Doré, 1995), we suggest that professional actors should be used, who are coached by a professional director and who use Stanislavski (1980) or method acting techniques that involve role taking, personal memories, and empathy, because this technique is likely to increase the believability of the portrayals. The concept is not to simply ask an actor to produce an expression, but to ask an actor to produce an authentic emotion that will automatically carry an expressive component.

Much EE work has always emphasized, and rightly so, pull effects and, in particular, cue utilization by observers. For this purpose, it is essential to create a sufficient number of conditions, repetitions, and controls, as well as a high degree of standardization. In consequence, we strongly reassert the utility of actor portrayals /see also Bänziger & Scherer, 2007) in studying EE and we present a new corpus constructed according to the theoretical and methodological desiderata outlined earlier and based on our past experience with similar attempts (Bänziger & Scherer, this volume).

Conclusions

Our intention in this chapter was to deliver three central messages. First, on a theoretical level, we wanted to clarify what the aim of EE research is. Should it really be the analysis of the spontaneous, sincere, unregulated expression of ongoing emotion processes in naturalistic settings? We argue (a) that current research procedures are incapable of obtaining

access to such expression instances, and (b) that these are a rare species indeed. Instances of pure push expressions occur very infrequently in adult members of modern society in which expressions are almost always determined by an interaction or combination of push and pull effects. We have suggested that it may be very useful to engage in a systematic examination of pull effects, including the iconic representations of specific emotions by patterns of expressive features, which probably account for a sizeable portion of EEs encountered in everyday life. We hold that for this purpose, actor portrayals provide a royal road.

It is obvious that actor portrayals cannot be treated as expressions of spontaneously occurring emotions. Actors pose an expression and although they may not be devoid of emotion when the use Stanislavski techniques to conjure up appropriate memories or images, the purpose of posing is clear – provide a plausible and believable impression of the emotion in observers. And the purpose of the research is obviously not to examine the nature of spontaneous emotion expressions or their underlying production mechanisms with the help of actor portrayals (although they can help to develop hypotheses, see Scherer & Ellgring, 2007). The aim is to examine a shared code of emotional signaling in social communication and to determine the way in which different types of facial, vocal, and bodily cues are used in signaling and in interpreting these signals. In this sense, systematic corpora of well designed actor portrayals can play a major role in examining the neurological and psychological processes that occur when we encounter certain patterns of multimodal emotion expressions. It seems to us that such corpora can also be of immense value for affective computing research and implementation, feeding the learning process described by Cowie et al. (this volume) and serving as appropriate criteria to evaluate automatic detection algorithms and synthetic expressions of emotional competent autonomous agents. It is to be hoped that the fruitless debate about what is natural and spontaneous in emotional expression and the rigid rejection of certain research paradigms, many of which have, like actor portrayals, shown

their utility in EE research, can be rapidly overcome. Given our evanescent object of study,

we need to use every possible angle of approach that is feasible.

Figure Captions

Figure 1. Adaptation of the Brunswikian lens model to Emotion Expression.