Author: Noordewier, Marret K.
Title: The dynamics of surprise and curiosity
Issue Date: 2016-02-24

CHAPTER 1

The temporal dynamics of surprise

A revised version of this chapter can be found in: Noordewier, M. K., Topolinski, S., & Van Dijk, E. (2016). The temporal dynamics of surprise. *Social and Personality Psychology Compass, in press.*

The Temporal Dynamics of Surprise

People are surprised when they are confronted with unexpected or schema-discrepant stimuli (Ekman, 2003; Horstmann, 2006; Meyer, Niepel, Rudolph, & Schützwohl, 1991; Meyer, Reisenzein, & Schützwohl, 1997; Smedslund, 1990). According to this view, surprise is a function of the difference between an expected and an actual outcome, like winning when one expected to lose (see also Teigen & Keren, 2003). This does not necessarily imply that people need explicit expectancies to be surprised. Surprise can also be evoked when people are confronted with something that does not fit their general schematic representation of the situation. When, for instance, a fluffy bunny hops into one's office, this is not surprising in the sense that people explicitly expected something different. This bunny may come as a surprise because it is not prototypical of one's representation of the situation, in this case, a business setting (see also misexpected vs. unexpected in Macedo, Cardoso, Reisenzein, Lorini, & Castelfranchi, 2009). Surprise thus reflects the difficulty in integrating or merely processing a disfluent or inconsistent event in a given representation (Kahneman & Miller, 1986; Maguire, Maguire, & Keane, 2011; Miceli & Castelfranchi, 2015; Topolinski & Reber, 2010a) and it is the interrupted state in which people do not (yet) know what has happened (see also Meyer et al., 1991, 1997).

In the current review, we will integrate theories and empirical evidence on surprise guided by a temporal dynamics perspective on surprise and sense-making (see Meyer et al., 1991, 1997; Noordewier &

Breugelmans, 2013; Pezzo, 2003). We distinguish surprise as the *initial state* that represents a cognitive interruption as a result of inconsistency or disfluency from *subsequent states* after people cognitively master the presence of the stimulus. To understand the nature and consequences of surprise, it is thus key to take time into account. We will arrange surprise research onto this interruption-to-mastering timeline with the goal to distinguish surprise from its consequences and to explain (seemingly) contradicting findings. We do this by a) critically discussing methods used in surprise research and b) reviewing different models and findings on the cognitive and experiential content of surprise, where we also integrate conceptually similar research in which it is not typical to refer to surprise (e.g., cognitive consistency, processing fluency, and immediate cortical responses such as the P300).

The Time Course of Surprise: From Interruption to Mastering

The time course of surprise starts with the detection of something unexpected. When people are surprised, ongoing thoughts and activities are interrupted, and attention is automatically directed to the surprising stimulus (Darwin, 1872/1998; Horstmann, 2006, 2015; Meyer et al., 1991, 1997; Plutchik, 1980; Reisenzein, 2000b; Scherer, 2001). The stronger the unexpectedness (e.g., because of stronger expectancies/schema's) the more intense the surprise (Schützwohl, 1998). Once the surprising stimulus is cognitively mastered (Attardo, 1997; Forabasco, 1992; Suls, 1972; Topolinski, 2013) surprise dissipates and other affective states follow depending on the valence of the surprising event (Ekman, 2003; Noordewier & Breugelmans, 2013; Tomkins, 1984).

Thus, people first respond to the unexpectedness of an event, independent of its valence, and then responses unfold contingent on the valence of the event, such as joy for positive events and sadness or anger for negative events. As such, the temporal time course of surprise involves the unfolding of responses from initial interruption to cognitive mastering—similar to the sequential appraisal perspective on emotion (Scherer, 2001), where the appraisal of an event being sudden, unfamiliar, or unpredictable (i.e., novelty-check) *precedes* appraising the event in terms of pleasantness, goal conduciveness, or coping potential. Importantly, this means that surprise itself is different from the state that people experience immediately after making sense of the stimulus.

The notion that time is important for surprise and its consequences originates from Meyer et al.'s (1991, 1997) cognitive-evolutionary model of surprise (for reviews, see Macedo et al., 2009; Reisenzein, 2000b; Reisenzein, Meyer, & Niepel, 2012). This model identifies sequential processing steps and states that surprising events first elicit the appraisal of schema-discrepancy, followed by interruption of ongoing information processing and the reallocation of processing resources together with the feeling of surprise. The process ends with the analyses and evaluation of the event and, if necessary, the updating or revision of schemas. We build on this model and mainly focus on the interruption part where people feel surprise. This initial interruption and feeling of surprise is, in our view, the main experiential component of surprise (see also Miceli & Castelfranchi, 2015), and experientially different from the part that follows when people analyzed the event and master the outcome. So, where Meyer et al.'s model focuses on the *sequence of processing steps* to provide a complete overview of *all* consequences after being confronted with a schema-discrepant event (cf. Meyer et al., 1997), we follow their logic to *differentiate* surprise as the *initial interrupted state* from *subsequent states* that follow after people cognitively master the surprising stimulus. In the terminology of Meyer et al.'s model, this initial state thus represents the feeling of surprise, the interruption of processing and reallocation of processing resources and the very first part of the analysis of the event—as long as people do not yet master event. This follows the definition of Reisenzein et al. (2012), who stated that surprise is the "output signal of the unexpectedness detector".

By distinguishing surprise from its consequences, we aim to provide a clearer conceptualization of what surprise is and how it feels. Specifically, in what follows, we will take the time course of surprise to 1) point out methodological implications, 2) review initial cognitive consequences that support the notion that surprise is initially interrupting and independent of the nature of the stimulus, and 3) integrate apparently contradictory findings concerning the experiential content of surprise.

Implications of Temporal Dynamics for Methodology

As methods differ in terms of timing, so do conclusions that can be drawn from them. Below we review various methods in terms of whether they primarily focus on the initial or later consequences of surprise. Specifically, there are many methods that induce surprise in participants, which offer the possibility to focus on surprise *while it happens*. Importantly, however, not all dependent measures are equally suitable to tap into the initial interruption. Indeed, many measures that claim to assess surprise, merely assess the consequences of surprise at a later stage. Moreover, there are also methods that do not focus on surprise while it happens but rather focus on associations, interpretations, or memories. This limits the possibility to study the experience of surprise without confusing it with the state that follows it.

Inducing Surprise

Inducing surprise is typically done by presenting stimuli that are incongruent to representations brought to the empirical setting (cf. *general knowledge structures*, e.g., Bless, Clore, Schwarz, Golisano, Rabe, & Wölk, 1996), or by disconfirming an expectancy that has been created within the empirical setting itself. Surprise has, for instance, been induced by a change in the experimental setting like changing the voice of the experimenter (Scherer et al., 2004) or even by transforming a complete corridor into a new room (Schützwohl & Reisenzein, 2012). Explicitly violating expectancies is often done via repetition-change paradigms (i.e., a different target stimulus after repeated exposure to a stimulus), such as presenting big/green letters after a series of one small/grey letter (Schützwohl & Borgstedt, 2005), or by changing the location of the experimenter in a sequential peek-a-boo game with infants (Parrott & Gleitman, 1989). Other expectancy-based methods include expectancy/ stereotype-inconsistent person descriptions or interactions (e.g., a sporty elderly woman; Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; Mendes, Blascovich, Hunter, Lickel, & Jost, 2007; see also Gendolla & Koller, 2002), semantically unpredictable sentences (Whittlesea, 1993), good news after having read that bad news was more likely (Sheppard & McNulty, 2002), or success after predicting failure is more likely (McGraw, Mellers, & Ritov, 2004).

Measuring (correlates of) Surprise

While these diverse induction methods most likely all result in surprise, dependent measures differ in terms of whether they focus on the initial interruption or on the state after cognitive mastering. Measures that can capture the initial interruption include facial expressions (e.g., Reisenzein et al., 2006), coding the absence or presence of a surprised expression using for instance FACS (Ekman, Friesen, & Hager, 2002); facial electromyography (fEMG, e.g., Topolinski, Likowski, Weyers, & Strack, 2009; Topolinski & Strack, 2015) or expression coding software (e.g., FaceReader; see Chapter 3), sometimes including measures of pupil dilation and vocalizations (Reisenzein et al., 2006). Also, various physiological measures can measure initial correlates of surprise, including arousal (e.g., Niepel, 2001), physiological threat responses (Mendes et al., 2007), and central nervous correlates obtained with EEG or fMRI (e.g., Alexander & Brown, 2011; Egner, 2011).

Many dependent variables do, however, *not* tap into surprise while it happens, but instead focus on consequences of surprise after cognitive mastering. First of all, self-reported surprise (e.g., Maguire et al., 2011; Reisenzein, 2000a; Schützwohl & Borgstedt, 2005; Teigen & Keren, 2003) is often assessed when participants already made sense of the surprising target. While people might be able to report how intensely they experienced surprise, this measure does not necessarily reflect the original experience of surprise while it happened because it is likely to be influenced by a retrospective evaluation of the event (see also Schützwohl, 1998). Also, when people report that they were *not* surprised, it is possible that they actually were surprised but meanwhile have re-interpreted the situation, such as in hindsight bias (Hawkins & Hastie, 1990). Similarly, self-ratings of related affective states or evaluative judgments most likely reflect consequences of surprise rather than experiential correlates of surprise itself (e.g., Mellers et al., 1997; Valenzuela et al., 2010). Concluding, self-reports of surprise and other affective states assessed after cognitive mastering are likely affected by various mechanisms different from surprise itself.

Indirect Methods

Besides inducing surprise in participants with inconsistent, disfluent, or unexpected outcomes, scientific knowledge about surprise has also benefitted from relatively indirect methods. Indirect methods do not focus on surprise while it happens but rather focus on associations, interpretations, or memories.

For instance, in linguistic methods participants rated the extent to which emotions, including surprise, fit words related to arousal and valence (Russell, 1980) or affective states, appraisals, and actions (Fontaine, Scherer, Roesch, & Ellsworth, 2007). While these methods are suitable to study the dimensional structure of longer lasting emotions, they may not reflect the actual experiential properties of the short-lived experience of surprise. The associations people report are potentially mixed and may not represent the initial short-lived surprise alone. The same holds for autobiographical recall of an unexpected/ surprising event. While recall supposedly regenerates feelings and participants appear able to differentiate between initial and later responses when asked to do so (Noordewier & Breugelmans, 2013), it is still possible that recollections of especially the initial interruption are influenced by the fact that participants recalled the whole experience, including the part that they already made sense of (cf. the peak-and-end rule of retrospective assessments, Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993, where the end phase of an experience dominates the later assessment).

Also, in some facial expression research, participants or actors are asked to express a certain emotion, based on an emotion label or a description of an emotional state (e.g., Du, Tao, & Martinez, 2014). This task requires actors to interpret the label or the description and come up with a fitting expression. These methods can be very suitable for perception and recognition studies (e.g., Ekman & Friesen, 1971), but are unlikely to cover the time course of surprise and particularly, the initial interruption phase of surprise. Moreover, when descriptions include the meaning of the surprising situation (e.g., happily surprised, like "receiving wonderful, unexpected news" or angrily surprised, like "when a person does something unexpectedly wrong to you", see Du et al., 2014; p. 7), the expressions probably include these interpretations rather the initial interruption, the state in which people experience difficulty integrating or processing an event.

Conclusion

The implication of the current temporal dynamics perspective for methodology is that research questions related to surprise while it happens (i.e., initial interruption) are best answered by methods that induce surprise directly *and* measure correlates immediately, whereas research questions related to the later consequences of surprise (i.e., once cognitively mastered) are best answered by methods and measures that assess consequences once people understood what has happened. With this perspective in mind, we will now discuss cognitive and experiential consequences of surprise.

Cognitive Correlates and Consequences of Initial Interruption

Several lines of research on neural mechanisms have addressed the immediate brain responses to schema-inconsistent or unexpected events. While in those approaches the term "surprise" is rarely mentioned, they examine the cortical responses to outcomes that conflict with predicted or anticipated outcomes (e.g., Bush, Luu, & Posner, 2000; Donchin, 1981), which renders this research as highly pertinent to the present review. As we will see, these instant brain responses seem to aid increasing processing depth to prepare cognitive mastering. Importantly, these responses are initially independent of the nature of the surprising stimulus, supporting the view that surprise is similar for all kinds of stimuli.

First of all, *conflict detection* (Bush et al., 2000) during *conflict monitoring* (Botvinick, Braver, Barch, Carter, & Cohen, 2001) can be conceived of as a form of surprise. In conflict detection, the medial prefrontal cortex, and particularly the anterior cingulate cortex (ACC) continuously monitor the contingencies of actions/decisions and their consequences and establishes expectancies of these contingencies from past to future actions and decisions (for reviews, see Alexander & Brown, 2010; Behrens, Woolrich, Walton, & Rushworth, 2007). If these contingency-schemata are violated, that is, if unexpected outcomes are encountered (either positive, neutral, or negative, e.g., Hajcak, 2012; Hajcak, Holyroyd, Moser, & Simons, 2005), the ACC shows an immediate increase in activity, as reflected by a negative deflection in the ongoing electroencephalogram (EEG), occurring less than 100 ms after the conflict arises, the so-called *error-related negativity* (ERN; for reviews, see Holroyd & Coles, 2002; Yeung, Botvinick, & Cohen, 2004). This ACC activity thus reflects a detection of violated expectancies within a very brief time frame and has recently been related to an aversive defense response (Hajcak, 2012; Hajcak & Foti, 2008). Functionally, ACC activity has been related to increased awareness (Lane et al., 1998) and the recruitment of control processes (Bush et al., 2000).

Another immediate cortical response to unexpected stimuli is the family of *P300* responses, or *novelty P3* (Donchin, 1981; Ferrari, 2010; Goldstein, Spencer, & Donchin, 2002). In a typical paradigm on this, participants receive numerous (often neutral) stimuli, many of which are frequent or expected, but some are infrequent or unexpected (e.g., in an oddball-paradigm, Ferrari, 2010; Picton, 1992). During this, time-locked signals in the EEG, event-related potentials (ERPs) are measured. This feature of expectedness as crucial manipulation in all these paradigms again renders this research as conceptually important for the more general psychology of surprise. It should be noted, however, that surprising stimuli are not necessarily the same as infrequent stimuli (see also surprise vs. rarity in Horstmann, 2015). Infrequent stimuli are only occasionally, but typically more than once, presented in the context of frequent stimuli (e.g., in an oddball-paradigm). As such, they can be expected or even awaited. Yet, because infrequent stimuli are also considered as interrupting, this research might help to understand surprise, as processes that are similar to surprise are involved.

Taken together, a variety of studies show for surprising compared to common stimuli a positive deflection in voltage in parietal ERPs with a peak around 250 to 500 ms after stimulus-onset, the P300 (e.g., Escera, Alho, Winkler, & Näätänen, 1998; Ferrari, 2010; Goldstein et al., 2002). Crucially, the more unexpected or seemingly improbable an experimental event is, the larger the P300 is (Donchin, 1981). Similar to error-related ACC activity, functionally, the P300 has been related to the recruitment of executive functions (e.g., Näätänen, 1990).

Finally, these immediate brain responses not only trigger more elaborate processing, they also elicit an involuntary attention allocation to the surprising stimulus (Escera et al., 1998; Horstmann, 2015).

Attention Allocation. To identify the cause of the surprise, people attend to the surprising stimulus in order to makes sense of it (e.g., Escera et al., 1998; Horstmann, 2015; Meyer et al., 1991, 1997; for a discussion on surprise vs. the orienting response, see Reisenzein et al., 2012). So, after the perception of a surprising stimulus, ongoing processing is interrupted and processing resources are reallocated to attend to the surprise (for a review on the surprise-attention link, see Horstmann, 2015). Supporting the temporal dynamics perspective, attending to a surprising stimulus takes time: Because the perceptual system is not prepared for the surprising stimulus, people first have to detect the discrepancy and shift their attention, which takes more time than attending to expected stimuli (Horstmann, 2015). Once people attend to the surprising stimulus (measured by for instance reaction times, Horstmann, 2002, 2005; or eye-tracking, Horstmann & Herwig, 2015) they subsequently gaze at it (presumably to make sense of it), which lowers attention to the non-surprising surroundings (i.e., surprise-induced blindness, cf. Asplund, Todd, Snyder, Gilbert, & Marois, 2010).

Attending to the surprising stimulus is a precondition for sense-making. Processing is then aimed at establishing whether the stimulus has any relevance and whether people can or should act (Meyer et al., 1991, 1997). This process delays responses. When participants are, for example, asked to respond as quickly and accurately as possible to the position of a series of circles in a repetition-change paradigm, they respond slower when a circle changes its color. Interestingly, this effect is most pronounced when the surprising stimulus is action-relevant (i.e., a circle) as compared to action-irrelevant (i.e., a diamond; Meyer et al., 1997). So, surprise is interrupting (Meyer et al., 1991, 1997), and particularly so when it is part of goal directed behavior. Also, whereas initial responses to surprise are independent on the nature of the surprising stimulus, during the action-relevance check people start showing differentiation in responses depending on the relevance of the target.

In line with this relevance view, surprise also seems to facilitate threat detection (Schützwohl & Borgstedt, 2005). Based on the notion that threatening aspects of surprise might require immediate action, Schützwohl and Borgstedt (2005) tested whether reaction times differed depending on the valence of words. In one study, participants were asked to respond to the location of a dot above or below two animal or sports words. After a series of trials, people were shown an unpleasant (e.g., spider) or pleasant word (e.g., seal) that was either surprising or not (color was [un]expected). While in the surprise condition the pleasant and unpleasant words were evaluated to be equally surprising, participants in the surprise condition responded slower to unpleasant words (e.g., spider) than pleasant words (e.g., seal). This delay in response to the dot was assumed to reflect an increase of attentive resources to the picture and as such explained as support for threat detection.

Conclusion

Initial cognitive consequences occur very fast after a surprising event and increase depth of processing for this event. This enhanced encoding and elaboration together with the additionally recruited executive functioning all aid the process of cognitive mastering (Pezzo, 2003; Topolinski & Reber, 2010b).[2]

[2] All this results in later increased memory for surprising stimuli (e.g., Sherman et al., 2012).

Crucially, these initial consequences are independent of the nature of the surprising stimulus. For instance, the novelty P300 occurs for surprising abstract neutral stimuli (Goldstein et al., 2002), acoustic material (e.g., Escera et al., 1998), and pictures (e.g., Ferrari, 2010). Also, it occurs for positive, neutral, and negative stimuli (e.g., Cano, Class, & Polich, 2009; Conroy & Polich, 2007; Ito, Larsen, Smith, & Cacioppo, 1998), and it is generally described as being insensitive to stimulus valence (Hajcak et al., 2005; although some find a more pronounced P300 for negative stimuli; Ito et al., 1998, Experiment 1). Only after some cognitive integration, people show differences in attention depending on the nature of the stimulus. In sum, cognitive consequences and correlates of surprise are initially the same for all kinds of stimuli. Over time—along with cognitive mastering—responses depend on the nature of the stimulus.

Experiential Content and Valence

The present temporal dynamics perspective of surprise is also informative for the ongoing debate on the experiential nature of surprise. Surprise "feels in a particular way" (Reisenzein, 2000b; p. 262), yet there is no clear answer to *how* it exactly feels to be surprised. Researchers seem to agree that surprise induces arousal, but at the same time, there are different perspectives on whether surprise feels good, bad, or neutral.

In work on the dimensional structure of emotions, people have been found to associate surprise with arousal (e.g., Fontaine et al., 2007; or astonishment in Russel, 1980) such as increased heartbeat and breath frequency, being restless and warm (for the complete list, see Fontaine et al., 2007). Direct evidence has been obtained by studies on the orienting response, where, for instance, skin conductance levels increased in response to surprising change in stimuli (e.g., Siddle, 1991; see however Niepel, 2001, who suggested that it is change rather than unexpectedness that drives arousal responses). In addition, people also appeared to have increased levels of arousal when they were instructed to mimic and experience surprise shown on a target picture (Collet, Vernet-Maury, Delhomme, & Dittmar, 1997).

It thus seems clear that surprise induces arousal, but there is less agreement on whether this is a pleasant or unpleasant state. Studies on surprise that included valence measures produced mixed results. According to some, surprise feels good (e.g., Fontaine et al., 2007; Valenzuela et al., 2010), according to others, surprise feels bad (e.g., Miceli & Castelfranchi, 2015; Noordewier & Breugelmans, 2013; Topolinski & Reber, 2010a), and it has also been argued that surprise does not have any valence (e.g., Mellers et al., 2013; Reisenzein & Meyer, 2009; Reisenzein, Meyer, & Niepel, 2012; Russel, 1980). The current timing perspective can explain and integrate these seemingly opposing findings (for a similar reasoning, see Noordewier & Breugelmans, 2013).

When zooming in on surprise as *initial interruption*, there are several reasons to depict surprise as having a negative valence. Surprise signals that an event has not been anticipated, which feels bad because it is in conflict with people's "need for prediction" (cf. Miceli & Castelfranchi, 2002, 2015). Specifically, for cognitive functioning, people need to anticipate events and validate these anticipations by facts; and if this does not happen, people experience distress (Miceli & Castelfranchi, 2015). In line with this, a prediction-error signal (i.e., ERN) has been related to an aversive defense response (Hajcak, 2012; Hajcak & Foti, 2008). Moreover, many studies in the domain of cognitive consistency (Gawronski & Strack, 2012), fluency (Topolinski, 2013; Topolinski & Strack, 2009a), and need for control (Kay et al., 2009; Rutjens et al., 2013) show that inconsistencies, disruption, and lack of structure are experienced as threatening and unpleasant as they represent a conflict with people's need for a predictable and comprehensible world (for overviews see Abelson et al., 1968; Gawronski & Strack, 2012; Miceli & Castelfranchi, 2015). Surprise implies a prediction error and reflects difficulty in encoding and semantically integrating a stimulus (Topolinski, 2012). When being surprised, reality thus does not make sense for a while. This state can be seen as dissonant, unstructured, and meaningless; all aspects that have been related to discomfort and threat (Elliot & Devine, 1994; Mendes et al., 2007; Noordewier & Breugelmans, 2013; Proulx, Inzlicht, & Harmon-Jones, 2012). Based on these longstanding research traditions, which developed relatively separately from literature on surprise, it thus seems plausible that the initial experiential disruption of surprise entails a phasic negative affect due to inconsistency or disfluency (cf. Alter, Oppenheimer, Epley, & Eyre, 2007; Winkielman & Cacioppo, 2001; Topolinski et al., 2009; Topolinski & Deutsch, 2012, 2013).

Support for this perspective can be found in a study by Mendes and colleagues (2007), who related expectancy violation to physiological threat responses. A threat (vs. challenge) response is associated with evaluating the situation as exceeding personal resources, which can be measured with distinct patterns of cardiovascular reactivity (Blascovich & Mendes, 2000). Mendes and colleagues reasoned that expectancy violation is threatening because it triggers surprise and uncertainty. In their experiments, participants interacted with a confederate who was a typical or an atypical member of his/her social category (e.g., white/Latino with high/low socio-economic status). Results showed that stereotype-consistent interaction partners were preferred over inconsistent ones (see also Schubert, Topolinski, & Strack, 2014; Swann, 1990) and atypical members elicited physiological threat responses (Mendes et al., 2007). Although further research is necessary to generalize this evidence to other surprising events or stimuli, it suggests that surprise indeed has threatening properties.

In addition, support for the reasoning that initial negativity unfolds to other affective states has been obtained by coding facial expressions over time from people who were positively surprised in TV-shows (Noordewier & Breugelmans, 2013). Participants were asked to indicate the extent to which a person in a picture felt negative or positive. The pictures were screenshots taken at one-second intervals after people were confronted with a positive surprise. Responses unfolded from relatively negative to positive, assuming to reflect the unfolding of initial responses to the unexpectedness to responses to the actual outcome.

In line with the view that *after cognitive mastering*, people respond to the nature of the stimulus, surprise has been found to be neutral or slightly positive when using linguistic methods (e.g., Fontaine et al., 2008; Russel, 1980). As discussed above, these results may reflect a mixture of surprise and associations of what followed it. Similarly, studies that included valence-related measures often administer them at a point in time where people already made sense of the surprising event. So, in terms of our perspective, they ask people how they feel or how they evaluate a surprising target *after* cognitive mastering. Then, people report being happy with a surprise gift (Valenzuela et al., 2010), satisfied by a better-than-expected performance of a product (Oliver, 1997), or delighted by an unexpected profit (Mellers et al., 1997).[3] Associations of the positive events with the word surprise might also explain why this word has more positive connotations than the word unexpected (for evidence suggesting this positive word-bias, see Noordewier & Breugelmans, 2013).

Conclusion

While more direct evidence might be needed to draw definite conclusions, the reviewed findings suggest that surprise has a negative valence, while the experience after cognitive mastering can be rewarding and positive (similar to the positive feeling of insight and punch line comprehension, Topolinski, 2013; Topolinski & Reber, 2010a). Thus, responses unfold from surprise to different states after people made sense of the stimulus. Only after cognitive mastering, people might feel good in the case of positive surprises. This dynamic

[3] Surprise can also amplify these responses (e.g., Mellers et al., 1997; Oliver, 1997; Valenzuela et al. 2010), due to an intensifying effect of residual arousal (Schachter & Singer, 1962), or because of comparison between expectancy and outcome resulting in contrast effects (e.g., Biernat, 2005).

was shown by the unfolding of facial expressions in the reviewed TV-show study (Noordewier & Breugelmans, 2013). Facial expressions thus seem to be particularly suited to study the temporal dynamics of surprise. It should be noted, however, that coding in the TV-show study focused on overall valence ratings rather than the distinct surprise expression or other specific facial features. When looking more closely to the facial expression of surprise, results become less clear.

The Face: Raised Eyebrows, Freezing, Frowning, or Nothing at All

The surprise expression has classically been conceived as involving raised eyebrows, widened eyes, and an open mouth/jaw drop (Darwin, 1872/1998; Ekman et al., 2002). Interestingly, the occurrence of this expression seems rare: Only a minority of surprised people actually shows this "typical" facial expression and those who do, show it only partly (i.e., one of the three components; e.g., Reisenzein et al., 2006).

For instance, studies with infants could not provide support for a clear facial expression of surprise. Instead, in response to surprises, infants show facial and behavioral freezing (Camras et al., 2002; Scherer et al., 2004). In these studies, a repeatedly shown toy was hidden and a different toy was revealed (Camras et al., 2002) or the voice of the experimenter changed via a speaker (see Scherer et al., 2004). Compared to baseline, surprised infants looked longer at the "impossible event" and showed more freezing (i.e., behavioral/facial stilling; see Camras et al., 2002; Scherer et al., 2004 for culture and age effects). However, they did *not* show particular surprise expressions (full or partial; but see Bennett, Bedersky, & Lewis, 2002; for a more infant expression studies, see Reisenzein et al., 2006).

Also in adults, the surprise expression is rare. Reisenzein and colleagues (Reisenzein et al., 2006) systematically tested whether surprise resulted in a surprised facial expression. They varied extremity of surprise, sociality of the context, duration and complexity of surprise, the degree to which surprise conflicts with exploration tendencies, all in different surprising contexts (e.g., when confronting participants with unanticipated photographs of themselves, Reisenzein et al., 2006; or changing a complete corridor, Schützwohl & Reisenzein, 2012). Participants reported being surprised and a majority of them also reported that they thought they had displayed a clear facial expression of surprise. Yet, only a minority of the participants actually showed a surprised face and often only partly (4-25%; Reisenzein et al., 2006; 34% in Reisenzein, 2000a).

Interestingly, in the studies reported here, the faces were mainly coded in terms of absence or presence of one or more of the three components of a surprised face (see above). It is, however, not completely clear whether participants who did not show a surprised face showed a neutral face, freezing, or signs of other facial expressions after a surprise. Sometimes, frowns were observed (Reisenzein et al., 2006: 9%/13% in Studies 3 and 8; Schützwohl & Reisenzein, 2012: 21%), as well as smiles (2-86% in Reisenzein et al., 2006; 26-71% in Schützwohl & Reisenzein, 2012; 8.6-12.1% in Scherer et al., 2004) and interest (on average 3.49 on a 5-point scale; Camras et al., 2002). In the other studies, no information about alternative expressions is available.

So, people do not seem to show the prototypical surprise expression, but they sometimes exhibit other facial signs like frowns (see also Topolinski et al., 2009; Topolinski & Strack, 2015) or smiles.

While these studies did not include time as a factor (except Study 7 in Reisenzein et al., 2006; see below), the different expressions can be placed on the interruption-to-mastering timeline. That is, freezing, frowning, the (partial) surprise expression, or no expression[4] most likely precede the smiles. Evidence for this unfolding can be found in the fEMG data in Reisenzein et al. (2006; Study 7). Participants were confronted with an unanticipated photograph of themselves in a facial judgment task. A slight increase of corrugator activity was observed after this surprising event and after 1-3 seconds this was followed by an increase in zygomaticus activity. So, a frown preceded a smile, which is in line with the current interruption-to-mastering logic (see also Noordewier & Van Dijk, 2015; Topolinksi & Strack, 2015).

Yet, based on the reviewed studies, it is not completely clear what the expression after a surprise looks like, as freezing, frowning, (partial) surprise, and neutral expressions all have been observed. Future studies should take time into account when trying to differentiate between different expressions (Noordewier & Van Dijk, 2015). In addition, other factors, like individual differences in expressivity and/or emotion regulation (e.g., Gross & John, 1997) might also partly explain

[4] It is possible that a neutral expression actually reflects freezing. Especially in stills, this might be hard to differentiate. To test the possibility of freezing in adults, it seems necessary to analyze videos, maybe even including bodily codings (e.g., Roelofs, Hagenaars, & Stins, 2010).

variations in expression. Importantly though, what freezing, frowning, and the surprise expression seem to have in common is that they all fit the initial interruption phase of surprise (Meyer et al., 1997). That is, freezing can be seen as a response that allows people to wait for further information, like sometimes observed with fear (cf. Scherer et al., 2004). Frowning has been associated with mental effort (e.g., Van Boxtel & Jessurun, 1993) and seems to be related to sense-making. Finally, it has been argued that the surprise expression helps to increase the visual field, which facilitates perception to identify the cause of the surprise (Darwin 1872/1998, see also Schützwohl & Reisenzein, 2012).

Conclusion

The temporal dynamics of surprise thus seem visible in the face. Analyzing facial expressions seems a very promising and suitable method to capture the interruption-to-mastering time line. Yet, more work needs to be done to fully understand initial interruption expression.

General Conclusion

Taking a temporal dynamics perspective on surprise (see also Meyer et al., 1991, 1997), we distinguished surprise as initial interruption from subsequent states that occur after people cognitively master the presence of the stimulus. Initially, consequences are the result of inconsistency or disfluency. Cognitive correlates (e.g., ACC/P300 activation, attention) are independent of the nature of the surprising stimulus and they all seem to increase processing depth to prepare cognitive mastering. Surprise is likely to be experienced as negative, as people do not like inconsistency, disruption, and lack of structure. Only after people cognitively master the stimulus, affective responses become in line with the valence of the stimulus. The cognitive and experiential consequences of surprise thus seem to fit the interruption-to-mastering timeline. An important implication of this temporal dynamics perspective is that methods to study surprise should be chosen carefully. Research questions related to the initial interruption phase are best answered by methods that induce surprise directly *and* measure correlates immediately, whereas research questions related to the later consequences of surprise are best answered by methods and measures that assess consequences once people understood what has happened. In sum, the temporal dynamics perspective on surprise helps to understand and integrate the nature and consequences of surprise.