Synthese
DOI 10.1007/s11229-016-1030-3

S.I.: FUTURE OF SOCIAL COGNITION

Defending the liberal-content view of perceptual experience: direct social perception of emotions and person impressions

Albert Newen[1]

Received: 18 February 2015 / Accepted: 19 January 2016

Abstract The debate about direct perception encompasses different topics, one of which concerns the richness of the contents of perceptual experiences. Can we directly perceive only low-level properties, like edges, colors etc. (the sparse-content view), or can we perceive high-level properties and entities as well (the liberal-content view)? The aim of the paper is to defend the claim that the content of our perceptual experience can include emotions and also person impressions. Using these examples, an argument is developed to defend a liberal-content view for core examples of social cognition. This view is developed and contrasted with accounts which claim that in the case of registering another person's emotion while seeing them, we have to describe the relevant content not as the content of a perceptual *experience*, but of a perceptual *belief*. The paper defends the view that perceptual *experiences* can have a rich content yet remain separable from beliefs formed on the basis of the experience. How liberal and enriched the content of a perceptual experience is will depend upon the expertise a person has developed in the field. This is supported by the argument that perceptual experiences can be systematically enriched by perceiving affordances of objects, by pattern recognition or by top-down processes, as analyzed by processes of cognitive penetration or predictive coding.

Keywords Direct perception · Rich content · Liberal content · Emotion · Person impression · Cognitive penetration · Predictive coding · Pattern recognition

✉ Albert Newen
albert.newen@rub.de

[1] Institute of Philosophy II, Ruhr-Universität Bochum, Universitätsstraße 150a, 44780 Bochum, Germany

Published online: 20 February 2016

1 Introduction

The debate about direct perception encompasses different topics, one of which concerns the richness of the contents of perceptual experiences. While some argue that the contents of perceptual experiences are exclusively low-level properties like edges, ridges, colors and shapes (Tye 1995), in recent years many people have argued that the contents of perceptual experiences can involve high-level properties or entities. The most important candidates for high-level properties in the recent literature are causal relations (Butterfill 2009; Siegel 2009), intentions (Pacherie 2005), actions and agency (Gao et al. 2009; Rutherford and Kuhlmeier 2013), and natural and artificial kinds (Bayne 2009; Siegel 2010). We also find the suggestion that the content of a perceptual experience can be the meaning of a sentence. This is argued by Siegel (2006), who relies on the phenomenological difference between listening to the same sentence in a foreign language before and after learning that language. In the same way, a phenomenological perspective is often used to argue for the rich content of our perceptual experience in social cognition, e.g. in Gallagher (2008) and Zahavi (2011). These phenomenological observations are in need of additional arguments, since they alone do not allow us to exclude the alternative explanation that there is only sparse content in perceptual experience, but that our phenomenology is a product of the combination of sparse perceptual experience with rich perception-based beliefs. The aim of the present paper is to offer a systematic defense of the liberal-content-view of perceptual experiences for emotions, and to outline how this view can be generalized to account for the formation of person impressions. A consequence of the view I defend is that direct perception can be acknowledged as a strategy of epistemic access in understanding others, since perceiving persons can result in a perceptual experience ("percept" for short) with a rich content. On the other hand, its role in explaining our understanding of others should not be overestimated, for we actually rely on a multiplicity of epistemic strategies to understand others, including simulation, theory-based inferences, interaction, and direct perception (Newen 2015). Before defending the liberal-content view of perceptual experience, we need to clarify which aspects of direct perception should be distinguished.

2 Situating my thesis about direct perception

There are many aspects to the debate about direct perception, and the richness of the content of perceptual experience is only one of them. Let me briefly mention two other aspects, simply in order to clarify that they are not part of the following discussion. From an ontological perspective, one question concerns the exact nature of the direct relatum of a perceptual experience: do I perceive an object in the world directly or only indirectly, where in the latter case the direct relatum is a sense-datum? It is a presupposition of the discussion that in the case of veridical perception we are directly related to objects in the world (metaphysical direct perception). Concerning informational processing, another question is whether the processes leading to a visual experience involve inferences or not. I remain agnostic on this, because the central question of

how rich the content of a perceptual experience is can be discussed independently of any position concerning the role of inferences.

To answer the central question of how liberal the content of a perceptual experience is, we first must establish a difference between perceptual experience and perception-based belief. If the difference vanishes, the claim that perceptual content can be liberal becomes uninteresting, because it no longer involves a claim about a perceptual state, but is rather a claim about a mixed state that integrates perception and belief. In that case, it is trivially true that we can deal with a liberal content based on perception. Thus, it is necessary to argue for the distinction. After doing so, I aim to defend a liberal-content view in the case of social perception, i.e. that content can involve emotions and person impressions. The arguments for this claim include (i) the central evolutionary role of social perception, (ii) the pattern theory of observable social cognitive phenomena, and (iii) evidence from pattern recognition and predictive coding that enables us to understand how social perception is systematically enriched.

3 The distinction between perceptual experience and perception-based belief

Why should we separate perceptual experience from perception-based belief? The alternative view is that they are so strongly intertwined that they are not distinguishable at all. Such a position can be maintained, for instance, on the basis of dynamic system theories that aim to give up any separation of perception, action, and cognition, including beliefs (van Gelder and Port 1995; Hurley 1998). I think this an implausible aspect of some variants of dynamic systems approaches, however; instead, I defend the separation by summarizing some standard arguments and adding some new considerations of my own.

One important observation is the stability of visual illusions, e.g. in the case of the Müller-Lyer arrows, one arrow looks shorter) even if we form the correct belief that both arrows have the same length (based on measurement). Further arguments for the distinction between perceptual experience and perception-based belief rely on specifying their different properties. We can certainly find such properties if we understand them as characteristic properties for standard cases, but I grant that those properties are not able to function as definitional criteria. However, this is not needed in order to argue for the separation of paradigmatic cases of perceptual experience, e.g. between a veridical visual experience and a belief based on that experience. To a first approximation, we can maintain this distinction by presupposing for a typical belief that its representation is language-based, where this includes representations which are based on natural language as well as linguaform representation (as is usually presupposed in a language of thought, for example).

Furthermore, although my focus in the present discussion is on *visual* experience, I intend to generalize such that it is possible to include experiences based on other senses, as well as multisensory experiences. How can we characterize core features of veridical visual experience in contrast to beliefs, where the latter are here understood as language-based cognitive evaluations at a time point? In a typical case, such experience

is image-like, while typical (language-based) judgments or beliefs are proposition-like, i.e. they are a composition of concepts. The feature of being image-like can be spelled out as being iconic, and thus as not having a canonical decomposition, as is true for typical cases of (language-based) beliefs (Fodor 2008). Sometimes, the properties of being continuous/discrete (Maley 2011) and analog/digital (Dretske 1981) are also used to characterize the distinction, and to make the related observation that features of perceptual experiences have no borderline, not even a fuzzy one, while beliefs and the concepts involved need some borderlines in order to transfer some minimally specific information. The latter features are connected with two arguments often used to contrast perceptual experience with concepts. Here, I want to use these as arguments to characterize perception as different from belief, while leaving open the definition of concepts (but see Newen and Bartels 2007; Newen and Marchi 2016). The first is the so-called argument for the "richness" of perceptual experience. To distinguish this from my central question about the liberal/rich content of the perceptual experience, I propose to call this the argument for the plurality of features in a perceptual experience. A perceptual experience involves a great many features of a scene, e.g. when seeing a picture of twenty people who are all dressed differently, we actually see all these people in their specific attire and their spatial arrangement, while at the level of belief we are forced to abstract from much of the plurality of details that we instantly perceive. A belief (even a rather extended language-based belief) is in principle much more abstract in its information than a perceptual experience, with its great plurality of detail.[1]

The second argument, for the fine-grainedness of perceptual experience, aims to show that there is not only a lack of abstraction in perceptual experience as compared to belief, but also a principled limitation on the fineness of grain of belief as compared to perceptual experience. The latter involves shades of colors and a variety of noises, shapes, and motional dynamics, which we cannot account for through our beliefs or the concepts involved in them. The latter are not fine-grained enough to account for the details we register in our perceptual experience. For example, when seeing two shades of red simultaneously, I can distinguish them even if they are of a very closely matching shade. But if I see them one after the other in different situations, I can no longer distinguish them, due to lacking a direct comparison. Since the concepts on which our beliefs are based have the core function of being useable for reidentification of the relevant properties involved in different situations (Newen and Bartels 2007), and reidentification in different situations is not epistemically possible for two closely related shades of red, it follows, in principle, that at the level of concepts (and beliefs involving them), we never reach the fineness of grain available to us when perceiving a scene. Another important criterion, characteristic of veridical perceptual experience, is a rather strict variation dependence on the actual sensory stimulus, while a (language-based) belief is typically not disposed to vary so strongly on a change in input stimuli. Thus, we have at least the following characteristic properties of perceptual experiences, which normally do not hold for beliefs, at least when the latter are language-based. Perceptual experiences are iconic, continuous and analog; they involve a plurality of

[1] My account of the higher-level abstractness of beliefs does not imply that I think of beliefs as combinations of abstract symbols in a language of thought, like Fodor (1975), but it remains consistent with beliefs being anchored in perception (for detailed view on concepts, see Newen and Marchi 2016).

features (lots of details), have a high fineness of grain, and co-vary rather strictly[2] with the actual sensory stimuli. This allows us to distinguish perceptual experiences rather clearly from the beliefs based upon them, as long as the beliefs are understood as language-based cognitive evaluations at a time point (it does not matter whether we are dealing with a natural language or a language of thought, as proposed by Fodor 1975). The central question is: Can perceptual experiences in a situation in which we can clearly distinguish a language-based token belief have a liberal content?

We need to be aware that there are other accounts of the concept of belief which I cannot address in this article. Let me briefly mention two central accounts, however, as well as the general lines of argument they suggest concerning the central question of how liberal the content of a perceptual experience is. Beliefs can be understood as dispositions (Schwitzgebel 2002), or as token mental states which may be realized in non-linguistic animals and children (Allen 1999; Newen and Bartels 2007; Glock 2010). In both cases it is much easier to argue that the content of the perceptual experience is enriched by the background beliefs. If one accepts that some beliefs are dispositions, then for those beliefs we lose the clear distinction between perceptual experience and perception-based belief which is a presupposition of the debate. A change of belief is in such cases characterizable as a change in the long-term disposition to behave. There is evidence that long-term dispositions are interwoven with the process of constructing a perceptual experience. This inspired the radical claim of Noë (2004) that without sensorimotor contingencies, which are simply dispositions, we are not able to develop our three-dimensional perceptual experience. If some beliefs are dispositions and these dispositions are a constitutive part of the process forming a perceptual experience, then beliefs can be a constitutive part of perceptual experience. The only way to block this conclusion would be to argue that those dispositions which are a constitutive part of the process forming a perceptual experience are radically different from those dispositions which are candidates for beliefs. This difficult challenge to argue for a clear difference of two types of dispositions has to be met by those who want to defend a principled claim that the content of perceptual experience can only be sparse and allow for some beliefs to be dispositions. Let us now glance at the account of beliefs as token mental states: possessed by non-linguistic animals and children. For those it is plausible to argue that such phenomena should better be described as "belief-like states" which can have a motivational role for goal-directed action. Like standard language-based beliefs of adults based on a perceptual experience they strictly co-vary with an ongoing perceptual experience, but they do not involve any linguistic concepts because non-linguistic animals and children do not have such things. Thus, the integration process leading to a perceptual experience in these subjects is an integration of the perceptual inputs and all relevant memorized perceptual images. Since there are no linguistic representations available for these subjects to be separated from the perceptual representations, it seems rather difficult to understand how the brain can prevent any relevant information available in the context from being integrated into the perceptual experience formed in the situation. Thus, if

[2] Exceptions are variations which are accounted for by phenomena like change-blindness or inattentional blindness (Noë 2004). More important is the role of imagination, which shares all the other features of perceptual experiences but lacks the co-variation with an actual stimulus.

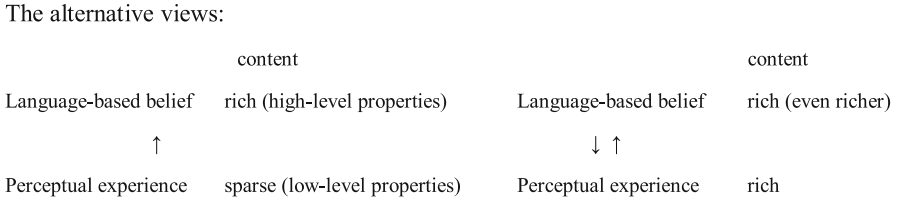

Fig. 1 The alternative views

non-linguistic animals and children have token beliefs at all, those are a central part of the process of forming a perceptual experience. The challenge for proponents of a sparse-content view is to show that despite the total lack of linguistic concepts, there is perceptual processing in non-linguistic animals and children which is clearly separable from further abstract processing, such that the abstract processing can deliver a belief which is different from the perceptual experience. If such a difference could be established, then additional argument would be needed to show that the content of the perceptual experience always remains sparse. These challenges for the two understandings of the concept of belief are worth noting, but cannot be further discussed here. In the rest of the article I focus on the most important and most challenging case for the liberal-content thesis: Can we defend the liberal-content view for perceptual experience in a case where such experience is clearly separable from any belief, seen as a language-based cognitive evaluation at a time-point? Thus, the question remains whether in such a case perceptual experience can have a rich content, or only a sparse one (Fig. 1).

A different way to pose the question is by asking whether perceptual experience can be enriched by concepts or language-based background beliefs or memorized images modifying the process of forming the perceptual experience, or whether the perceptual experience is a result only of rather rigid bottom-up processes always leading to a sparse content. If we can establish enrichment processes as part of percept formation, and show that there can be a systematic top-down influence of concepts, memorized images or language-based beliefs on the formation of perceptual experience, then this indicates that the contents of a perceptual experience can be rich, insofar as they can be enriched by these influencing beliefs, concepts or memorized images.

In arguing for the liberal-content view, I proceed by introducing some non-social examples which furnish evidence that perceptual experience has a rich content: to this end, I discuss affordances, and argue that perceptual experience can be enriched by memorized image contents, citing examples of conceptual contents of which one main case is the perceptions of chess experts as compared to novices (Sect. 4).

4 Evidence for a rich content of perception: organizing perceptual experiences by integrated patterns of non-social objects or situations

One possible strategy in arguing for a rich content for perceptual experiences is to cite evidence that the structuring of the perceptual input is essentially shaped by rich content that could not be understood as part of the perceptual experience if it were

always sparse. First, when we perceive objects, we not only perceive the object but also the possibilities involved in dealing with it, i.e. its affordances: when we perceive a cup, for instance, it affords us to reach out to it and grasp it by its handle, without any instruction to do so (given sufficient motivation and attention). Several studies indicate that merely viewing an affording object triggers motor activations in the brain (fMRI confirmation by Chao and Martin 2000; EEG-support by Proverbio et al. 2011). This can best be understood in terms of perceiving affordances (Gibson 1979), since the motor activations are correlated with a conscious registration of the possibility to act not only in human adults but also in non-linguistic animals and children.[3] Since the registration of affordances remains perceptual but includes the enriched content of the perception, namely possible action with the object, this is a first step beyond a sparse-content view.

A follow-up question is how complex these affordances can be. To foster my liberal-content view, I want to make use of evidence provided by modern developments in Gestalt psychology, since Gestalt principles organize sensory input and thereby can enrich the perceptual experience (in principle in contrast to a perceptual experience organized without Gestalt principles). It has been shown that this not only involves perceptual grouping principles like proximity, good continuation, closure, symmetry or parallelism (Wagemans et al. 2012a, b), but also that past experience plays a role in basic processes like figure-ground segregation (Trujillo et al. 2010). This raises the possibility that memorized higher-level information may enrich the content of the perceptual experience. To prove that this can actually be the case, we need convincing case studies: a well-known case is the picture of the "Dalmatian dog" which at first glance looks like an unorganized structure of black and white dots. If we receive the information that there is a dog in the picture then we quickly see it, and once we detect the dog the perceptual content changes: we can never again see this picture as one of random black and white dots. To provide an example of the influence of complex background memories, I elaborate below on the perceptual experience of chess masters looking at a chess board during a game, defending the claim that chess experts rely on complex visual pattern-matching. With her superior organization of knowledge, a chess expert can rapidly perceive a promising move, just like a medical expert can quickly notice an inconsistency in a suggested diagnosis. The process of using this information smoothly relies heavily on fine-grained pattern-discrimination and pattern-matching (Gobet 1997) in the relevant situation, rather than on drawing inferences or making judgments. This is supported (a) by general observations about the way people recall chess positions. When seeing a chess board that contains a real, meaningful arrangement, chess experts excel when compared to novices in recalling positions, but perform no better in the case of scrambled, impossible positions (Gobet and Simon 1996). This indicates that they are able to see meaningful patterns that a novice cannot. Furthermore, (b) we can rely on cognitive studies conducted in recent

[3] One might worry whether the motor activations in the brain are a constitutive part of the perceptual process or only a causal consequence of it. A conclusive account of this distinction is not possible here. But it is sufficient to highlight that in a context of registering an affordance of an apple, such as being edible, this registration is strongly intertwined with the act of perceiving the apple and that this act of registering the affordance can happen even for agents who are not able to form language-based beliefs.

years in order to rule out the alternative that the chess expert has the same perceptual experience as the novice and differs only in their judgments. We need evidence that the higher-level pattern-activation process used by chess experts is part of the perceptual process and belongs not only to judgment. We can find this in two steps: (i) In a recent neuroimaging study, the comparison of intact versus disturbed perception of global gestalt indicated the significant role of the temporo-parietal junction (TPJ) in the intact perception of global gestalt (Huberle and Karnath 2012). This location corresponded well with the areas known to be damaged or impaired in patients with simultanagnosia after stroke or due to neurodegenerative diseases. Simultanagnosia is clearly a deficit in the perceptual process, typically consisting in patients perceiving all the components of a complex object but not the pattern that the object constitutes: e.g. patients can see the vertical column of 8 Xs and the horizontal column (attached at the bottom right) consisting of a further 4 Xs. But they cannot see the L formed by these Xs. Of course, the patient can see Ls, and may look at the general shape of an X which is constituted out of Ls. In this case, the patient can see all the Ls but lack the perceptual experience of an X. The study concluded that the TPJ plays an important role in the integration of individual items in a holistic perceptual experience. (ii) To return to chess experts, Karnath's group (Rennig et al. 2013) started with the working hypothesis that chess experts who rely very heavily on holistic perceptual experiences should rely significantly more on TPJ activation when perceiving complex chess situations. The meta-analysis of the fMRI data from four studies that compared chess experts with chess novices during the presentation of complex chess-related visual stimuli proved their hypothesis. They observed higher TPJ signals in chess experts in comparison to novices during observations of complex chess positions. This difference was consistent across different tasks in several independent experiments (Rennig et al. 2013). I take this case study in expert perception to show that perceptual experience can be influenced by different cognitive processes relying on the expertise, where these processes enable holistic pattern formation to form the percept. This case study should be seen in the context of several studies on perception in sports activities, where strong differences between the perception of experts and novices were also reported (e.g. perceiving the direction and force of squash strokes, Abernethy et al. 2001). Thus, it seems that expert perception of objects or processes leads systematically to differences between the perceptual experiences of experts and novices, due to a knowledge-based enrichment of perceptual experiences in the former case.

Let me consider one alternative explanation: perhaps the expert knowledge does not enrich the perceptual experience but rather modifies a perceptual module which produces a different perceptual experience than before the modification. Since the central claim of this article concerns the richness of the content of our perceptual experience but not the mechanism of producing the richness of the content, this alternative interpretation is not a challenge as long as it allows for the content of the perceptual experience to be richer after expert knowledge has modified a perceptual module. Thus, if one accepts a modular view which still allows for some modifications of modular processes by expert knowledge (*weak* modularity view) and thereby allows for richer contents of visual experiences despite relying on a modularity view, then there is no conflict with the liberal-content view.

A conflict arises only if one combines a modular view of perceptual processing with the claim that perceptual processes are (always) impenetrable by concepts or expert knowledge (*strong* modularity view). Then it follows that if the only change from one situation to another is an activation of a concept or background knowledge, then the perceptual experience must remain the same sparse content and any impression of modified perception can only be interpreted as a change at the level of judgments based on the constant perceptual experience. This leads directly into the heart of the intense debate about cognitive penetration of perceptual experience (Macpherson 2012; Vetter and Newen 2014; Firestone and Scholl forthcoming). Just to be clear: there is no need to defend the possibility of cognitive penetration to make an argument for a liberal-content view because the liberal-content view does not imply a strong modularity view (with impenetrability of perceptual processes)[4]; but if we have evidence for the possibility of cognitive penetration then this is additional support for the liberal-content view because then we allow for two mechanisms which may produce rich contents of perceptual experiences, namely (long-term) modified perceptual modules or situational top-down influences by activations of concepts or background knowledge. Thus, I refer the reader to additional arguments for cognitive penetration (Macpherson 2012; Vetter and Newen 2014),[5] and I argue here for cognitive penetration only in the special case of emotion recognition, which is in focus for the central part of the paper.

Having discussed evidence for the liberal-content view in non-social contexts, I now switch to social cognition to deliver additional specific evidence.

5 Emotion recognition as a perceptual process leading to a perceptual experience with a rich content

To make a convincing case for emotions, I shall argue that emotion recognition in the case of basic emotions is a *perceptual* process since it combines central features, i.e. it is evolutionarily very important, very quickly formed, and action-guiding in relation

[4] See Toribio (2015) for a position combining impenetrability with a liberal-content account. But in her proposal it is not clear enough how the enrichment can be realized, especially since the enrichment has to be realized by a nonpenetrable module. Even in token cases without cognitive penetration, I allow for a weak modularity view.

[5] To strengthen the argument for the liberal-content view, I highlight one case study supporting the cognitive penetration of perceptual experience by conceptual knowledge: Winawer et al. (2007) demonstrated the influence of basic color concepts on perception. They presented Russian and English speakers with color swatches of different shades of blue. The experiment is based on different ways of categorizing shades of "blue" in both languages: Russians lexicalize the category blue with two basic level terms: "siniy"' for darker blues and "goluboy" for lighter blues, while the English have just one basic-level term "blue." The students were asked to decide as quickly as possible whether a top color exactly matched a color on its left or on its right. While all shades of colors were in the category "blue" for the English, the colors were part of the two different basic categories of blue for the Russian speakers. Since we expect to perceive the difference between different color categories more intensely than within-category differences, we should expect that the Russian speakers have quicker reaction times on between-category trials than within-category trials, while the English speakers should show no such effect. And this is exactly what was observed. Further evidence of a top-down influence of linguistic labels or categories on perception is reviewed by Lupyan (2012).

to online observations of typical features constituting the emotion[6] (even for non-linguistic children) (Sect. 5.1); then I argue that the content of emotion recognition is *rich* by showing that emotions are individuated as patterns with a rich content (Sect. 5.2). Furthermore, the richness of the content is supported by the observation that the process of emotion recognition is a process of pattern recognition such that memorized patterns can enrich parsimonious inputs (Sect. 5.3).

5.1 Emotion recognition as a perceptual process

Why is the process of recognizing a basic emotion like fear, anger, joy or sadness a perceptual process which leads to a perceptual experience and not always to a judgment?

(a) Evolutionary importance: It is extremely important for humans as hyper-social beings to quickly and reliably notice emotional states and the intentions to act of the human beings they interact with. This makes it very plausible that this information is part of our perceptual experience, especially since this is partially shared with some nonhuman animals who lack representations of beliefs in any linguistic format. For present purposes, I focus on recognizing emotions. The roots of this ability are anchored in nonhuman animals. The recognition of facial expression is known to be relevant for nonhuman primates (Leopold and Rhodes 2010). Chimpanzees and rhesus monkeys rely on the accurate interpretation of facial expressions since this—as with humans—is used for nonverbal communication in their social group. It has furthermore been shown that nonhuman primates do not simply rely on feature detection, but on a combination of detecting features and configural pattern (Parr 2008; Parr and Heintz 2009). Furthermore, there is a famous literature which argues for a psychoevolutionary account of human emotions (Griffiths 1997, 2004; Panksepp 1998, 2000). These authors bring together a lot of interesting evidence for the evolutionary basis of at least some emotions. The main evidence concerns e.g. neural circuits of fear, which are shared between humans and some nonhuman animals, as well as evidence for the existence of basic emotions. Ekman has shown that some emotions, including anger, fear, happiness, sadness, disgust, etc. are widely shared across all the cultures he has investigated, and that these emotions are easily and reliable recognized (Ekman and Friesen 1971).

(b) The *online* action-guiding role of a percept in the case of registering an emotion: For smooth interaction, we need to be able to register online the emotional situation of the interaction partner on the basis of social signals, independently of whether we in addition have language-based beliefs. This has been nicely demonstrated by recent work on naturalistic early parent–infant interaction with infants between 12 and 36 month of age. The interesting result was that synchronic interaction and

[6] The phrase "observing typical features constituting the emotion in this situation" is meant to be a neutral description of the fact that emotion recognition is based on a perceptual process of registering a person, her face, her body posture etc., but leaving open whether the content of this perceptual process is rich or sparse because this is what the argument is aiming for.

a smooth interactive dialog was shown not to be possible in a situation in which there was an established dyad of interaction with severe emotional neglect. The latter is supposed to be based on inadequate automatic social signal processing (Avril et al. 2014). The general argument is that it is plausible that a content (e.g. a rich cluster of social features) is part of a perceptual experience (not of a belief) if the percept is involved in the online shaping and transforming of our dispositions to act according to this content. I can perceive an apple being reachable, and I can perceive this very apple as ripe and as being edible (Nanay 2011). If we generalize this to emotions, I can perceive someone's anger because I am disposed (online) to defend myself. Since the online perception of another person's emotion modifies (online) our dispositions to interact according to the rich unit of features constituting the emotions, it is plausible that the content of the perceptual experience is the rich emotion itself.

These aspects indicate that the content of registering another's emotion (at least for basic emotions) is actually part of the perceptual experience, and not only of the belief formed on the basis of that experience (if such a belief is formed at all).

5.2 Emotions are individuated as rich patterns of characteristic features

Why should we accept that the content of an emotion is rich rather than sparse? We can start with the common observation that, from a first-person perspective, we often have a rich phenomenology while registering the emotion of another person (call this the argument from phenomenology). But this observation in isolation does not settle the matter, because this rich phenomenology could either be an aspect of the perceptual experience alone or a product of the association of the perceptual experience with some post-perceptually formed beliefs. But I can offer the following additional arguments to bolster the claim. I first argue that emotions are individuated as rich patterns of characteristic features (Sect. 5.2), then take up the question of whether such a rich pattern can be the content of the perceptual process that is involved in (online) emotion recognition (Sect. 5.3).

What exactly is the mental object that I want to claim is directly perceivable? I defend the view that we can perceive those mental phenomena which are individuated as a *pattern of characteristic features*. Let me illustrate this view by discussing the emotional episode of test anxiety (the basic emotion of fear towards an aggressive dog is discussed in Newen et al. 2015). We are using the expression "test anxiety" to speak about a type of fear that is individuated by a pattern of characteristic features. These include: (1) a typical physiological arousal that is a consequence of bodily changes due to changes in the autonomic nervous system, including increased heart rate and flat breathing; (2) a typical behavior or behavioral disposition, including flight or freezing behavior; (3) a typical facial expression, gesture, or body posture, etc.; (4) a typical phenomenal experience of fear; (5) a typical (explicit) cognitive evaluation of the forthcoming test (that it is important and that the subject will fail it). Furthermore, every emotional episode has (6) an intentional object, e.g. the upcoming test in mathematics. Features 1–5 are integrated into an (often implicit) appraisal of

the intentional object as dangerous. The appraisal is integrated with the intentional object into the unity of the emotional episode. This view allows that some features could be missing (e.g. the facial expression may be inhibited, due to intense training in "poker face"). However, as long as a minimum of features is realized, we still have an episode of test anxiety. Depending on the type of emotion, some features may be necessary, e.g., in the case of test anxiety, the cognitive evaluation of having a forthcoming test that one is not able to master. It also seems to be the case that the typical physiological arousal is not merely a typical but indeed a necessary component since any emotional episode involves some type of characteristic bodily arousal (although we can allow for some variation of the type of arousal involved in the same type of emotion, Barlassina and Newen 2013). One might wonder why I do not include neural correlates. Since I argue from a position of antecedent naturalism, neural correlates are not an extra component in addition to the characteristic features already mentioned above—we might mention neural correlates as an informative aspect for the individuation of certain features of emotion, but we do not have to, since they concern the same features that have already been mentioned, with information accessed in a different manner.[7]

This account of emotional episodes as a pattern of characteristic features comes with several advantages: in particular, it allows us to account for both evolutionarily anchored and widely shared basic emotions, as well as for culturally varying emotions (Newen et al. 2015). Basic emotions (Ekman 1972) as evolutionary anchored may be realized without any cognitive evaluation, e.g. seeing the dog is sufficient to produce the appraisal of fear directed towards the dog (Le Doux 1998). The pattern theory can account for this, as well as for culturally varying emotions of the same type, because the pattern remains the same even if one aspect is missing or developed in a slightly different manner. The pattern theory is furthermore able to account for cases in which there is not just cultural variation of an emotion type, but also when there is a radical cultural difference in the categories of emotion (see Welpinghus and Newen 2012), e.g. "song", on the island Ifaluk, is the word for a righteous form of anger, which differs from the concept typical in Western cultures in that it does not involve any aggression. This indicates that it is really a different type of emotion. These examples of rather different emotions motivate the social constructionist account of emotion, which takes emotions to be social constructs that are completely shaped by cultural conventions (Lutz 1986; Heelas 1986). The pattern theory can account for such cases, while accepting the evolutionary anchoring of basic emotions which unfold in ontogeny into "cognitive" emotions (Zinck and Newen 2008). Thus, the pattern theory can account for emotions as unities that are neither arbitrary conventional constructions, nor completely determined by evolutionary aspects. One way to spell out the ontology of a pattern theory of emotion is to argue that the pattern of characteristic features form a homeostatic property cluster, constituted by a combination of causal and social processes (Welpinghus 2015). However, a detailed ontology of emotions goes beyond the scope of the present article (Fig. 2).

[7] Neural correlates may be used in a clinical or scientific context to infer whether a person is experiencing emotions, but are not used in ordinary contexts since we cannot access them in ordinary interactions.

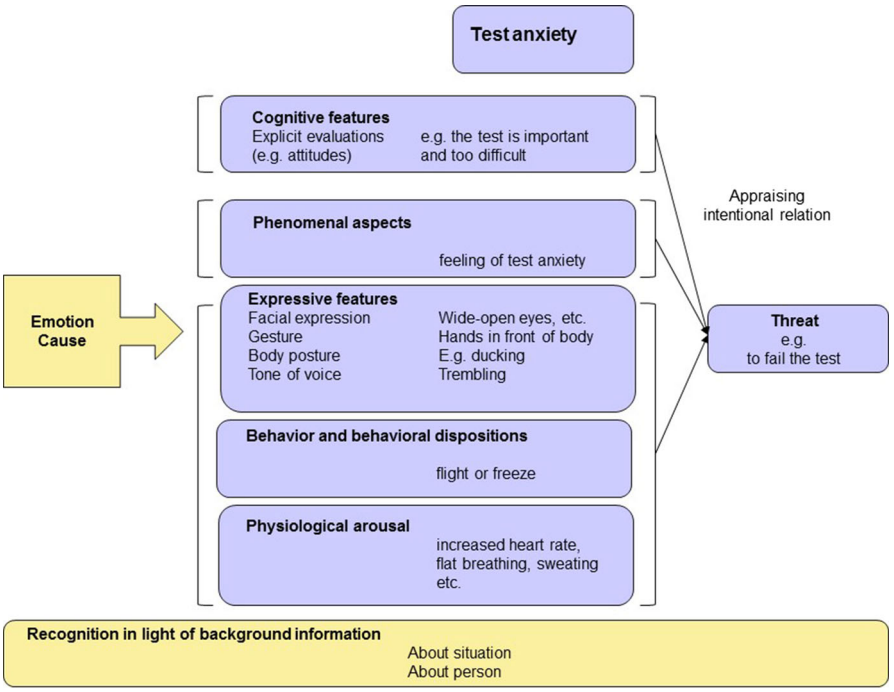

Fig. 2 The pattern of test anxiety

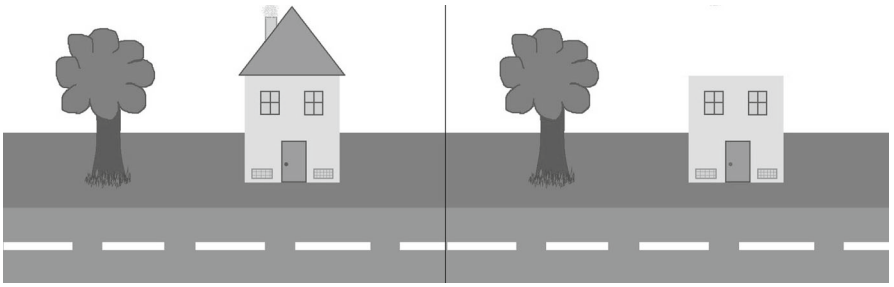

Fig. 3 Pattern recognition when seeing a house

5.3 Emotion recognition as pattern recognition: a process of enriching the content of perceptual experience

If it is accepted that emotions are individuated as patterns, we can provide a nice description of the process of emotion recognition as one of recognizing the pattern of the relevant emotion type. Before applying this to emotions, let me illustrate the core idea using the typical childhood example of drawing a house (Fig. 3).

We easily agree on the characteristic features of a village house when we focus on the features visible from the outside. These include walls, windows, a front door, a roof, and a chimney, as well as a garden with trees and bushes. These features form

a pattern of a house, even if in one drawing the front door is missing, or if in another there is neither a roof nor a chimney. The pattern can remain the same even if the components that realize it vary.[8] This can be transferred to the *recognition process*. I can recognize the drawing as one of a house, even if one or two of the characteristic features of the house are missing. How is this possible? Perceiving an object is not a purely passive process, like taking a photograph: it is a constructive process. All modern theories of perception account for this constructive component (e.g. O'Regan and Noë 2001 theory of enacted perception), as do theories of cognitive penetration (Macpherson 2012; Siegel 2012) and theories of predictive coding (Hohwy 2014). One important aspect of the constructive process is the enrichment of selected core sensory information. One way of realizing this is by the activation of a rich memorized mental image that best suits the core sensory information. If we have learned the relevant pattern of how a house looks from the outside and memorized a respective mental image, then seeing a child's drawing initiates a process of interaction between bottom-up and top-down processes that includes the activation of this stored mental image, such that it enriches the core sensory information to form a perceptual experience of seeing a drawing of a house.

The same process of pattern recognition takes place in the case of recognizing emotions (Newen et al. 2015). Let me elaborate upon this claim. The relevant pattern of test anxiety is formed either on the basis of having personally experienced the situation of taking an important test together with a deep fear of potential failure, or after having observed others in such situations. Usually, we can rely on both, at least in cases of basic emotion. In such a standard case of having undergone some text anxiety in the past, and having experienced several situations of other people in a very similar situation, I can optimize the mental pattern of test anxiety. If I have undergone the emotion myself, it will be easy to register some of the most relevant features constituting the emotion, e.g. the intentional object (the important upcoming or ongoing test), the typical cognitive evaluation and the behavioural disposition, as well as the typical phenomenal experience (see Fig. 3).

But how does this help to recognize a case of test anxiety when observing others? I cannot easily access the phenomenal experience of the other, or know whether she is undergoing the relevant cognitive evaluations. In addition to integrating the features constituting the emotion into a pattern, one usually learns to associate the type of context that is typical (e.g. a classroom situation or a person waiting in front of an office with a sign that reads: "Please do not disturb. Oral examinations"). Contextual information can be rather easily observed and can help to activate the relevant pattern. In the case of recognizing a house, it helps a lot to see a house (even if the drawing is bad) in the context of a garden with trees near a street, and if it is the right size relative to the surrounding elements. Analogously, in the recognition process we benefit from a direct coupling of the characteristic features of an emotion with paradigm contextual information. If I am the examiner, it is often sufficient to register the core components of the fearful face and trembling voice of a candidate to activate the mental pattern of test anxiety, and to develop the perceptual experience of the candidate with test anxi-

[8] This idea is also used by v. Savigny (1988) to develop an interpretation of the later Wittgenstein.

ety. From knowing the situation as an examiner, I can activate the relevant intentional object as part of activating the mental pattern of test anxiety. This enrichment realized through the process of pattern recognition can most plausibly be evaluated as an enrichment of the perceptual experience, in accordance with recent cognitive models of perception, e.g. Ernst and Bülthoff (2004). Furthermore, pattern recognition is one standard type of process involved in the construction of a percept which is accepted almost without dispute as a process that enables completing a percept by activating a memorized pattern on the basis of parsimonious inputs. There is a debate whether pattern recognition is mainly a process that relies on template matching, feature detection, or prototype evaluation, or some other mechanism but it is commonly accepted that pattern recognition is involved in face recognition and in visual Gestalt recognition of different complexities. Putting the three arguments together—(i) emotion recognition is a perceptual process, (ii) emotion recognition involves a rich content, i.e. an emotion pattern, and (iii) it is a process of pattern recognition—makes it very plausible that emotion recognition (at least of basic emotions) is a perceptual process resulting in a perceptual experience with a rich content.

In general, in addition to the core features of the pattern constituting a type of emotion, we learn to associate the typical cause of an emotion and the typical context. We are also able to integrate information about specific individuals and how their facial expression of emotion works, e.g. if a person in the family is suffering from Parkinson's disease and thus has strong constraints in moving their facial muscles, an expression of joy may be difficult to register at first; however, after a while one learns to see this person's joy by having stored a special mental image of his or her expected facial expression. As we have illustrated above (Sect. 5.1), there is a lot of evidence that perceptual learning processes enable us to have perceptual experiences with a rich content, e.g. the case of chess experts. Putting this evidence together allows us to conclude that the content of a perceptual experience can be rich, because the perceptual experience is systematically enriched by stored mental images. How rich that content is depends on the available or learned mental images, concepts or beliefs. Experts in a field can perceive much more than laypersons, but in all these cases there nevertheless remains a difference between a perceptual experience and a perception-based judgment (at least when the latter has a linguistic format).

6 Further processes of enriching the perceptual experience: cognitive penetration in emotion recognition

To complete the argument for the claim that emotion recognition is perceptual and has a rich content, I present a case study of emotion recognition which can best be understood as a case of cognitive penetration (Sect. 6.1) and I propose that predictive coding is a plausible mechanism for top-down influences that can produce the rich content of a perceptual experience (Sect. 6.2). If we take all three of these points together, it is very plausible to claim that the content of perceptual experience is rich in the case of emotion. This will also be outlined in the case of person impressions (Sect. 7).

6.1 A case study of perceiving emotions

Let me start by describing an impressive example that demonstrates how background knowledge of a situation penetrates emotion recognition. In an experiment by Carroll and Russell (1996), participants had to evaluate the emotion expressed by a human face. Subjects were presented with combinations of faces and situations. The target stimuli were still photographs of posed facial expressions, selected from among the prototypical facial expressions of fear, anger, disgust, joy, or sadness, collected in Matsumoto and Ekman (1988). These prototypical facial expressions have the peculiar characteristic of being reliably evaluated as expressing the same emotion across different subjects and cultures (Keltner et al. 2003), if no additional information is available. In the experiment, situations were provided in the form of short stories concerning the people depicted in the stimuli. Such stories were designed to trigger an emotional response of fear, anger, sadness, or disgust. Subjects were first told the story, and then shown the picture. They then had to evaluate the emotion expressed by the face by choosing one of the few emotion labels for basic emotions. Here is a (slightly condensed) example of one short story.

> A woman wanted to treat her sister to the most expensive, exclusive restaurant in their city. Months ahead, she made a reservation. When she and her sister arrived, they were told by the *maitre* that their table would be ready in 45 min. Still, an hour passed, and no table. Other groups arrived and were seated after a short wait. The woman went to the *maitre* and reminded him of her reservation. He said that he would do his best. Ten minutes later, two other couples arrived and were seated immediately. The woman went to the *maitre* who said that all the tables were now full, and that it might be another hour before anything was available (Carroll and Russell 1996, p. 208).

After having listened to the story, the participants were shown a paradigmatic facial expression of fear (according to the Ekman studies). The researchers found that when presented with such contextual information, the vast majority of subjects evaluated the face as expressing anger. When the contextual information was not presented, however, subjects evaluated the same face as expressing fear.

Of course, we need additional arguments to ensure that this effect demonstrates the influence of background knowledge on perceptual processes, and that it is not only a product of modifying our perception-based judgment. In its short version, the argument runs as follows: The isolated recognition of all Ekman faces for basic emotions is highly reliable and distinct (Ekman and Friesen 1971), i.e. people do not mix up faces of fear and anger. This means that for the different facial expressions of fear and anger, there is also a highly reliable coupling between perception of the face and the associated correct judgment. It is implausible to presuppose that under normal perceptual conditions (no ambiguous stimulus, normal perceptual abilities of the subject, good external conditions for perception, no learning of new concepts takes place), the vast majority of people suddenly substitute the correct coupling by a false one between the perceived face (which is still supposed to be a fearful face) and the judgment about the type of emotion (which is supposed to be anger).

Since the reported experiment shows that with the contextual background story, the standard Ekman face of fear is falsely judged to be a case of anger, the best explanation is to presuppose a change of both the perceptual experience and the judgment, which remains correctly coupled with the experience. The alternative would be that only a change of judgment takes place while the perceptual experience remains the same: but why should the judgment now be thought false in relation to the experience, despite there being normal perceptual conditions including an unambiguous stimulus (the face), high reliability in distinguishing fear and anger when seeing a facial expression, etc.? There is also no evidence that the test persons ignore the perceptual input, since the cognitive penetration of the emotion recognition by the contextual story is constrained: it does not work for a joyful face combined with a story of disgust or of sadness. Thus, given the constancy of normal perceptual conditions, we suggest that an incorrect perceptual experience is caused by the background story, while the highly reliable ability to recognize standard Ekman faces is not impaired.[9] Thus, the effect that the perceptual experience of a culturally universal facial expression of fear can be modified by a typical anger-story can best be explained as a case of the top-down influence of beliefs (introduced by the story). To add plausibility to this account, I offer an analysis in the framework of predictive coding.

6.2 Analysis in the framework of predictive coding

In recent approaches to predictive coding (e.g. Clark 2013), perception is presented as a process of enriching perceptual experience. Perception is a process that attempts to match incoming 'driving' signals with a cascade of top-down predictions that aim to cancel it out. It is seen as a basic and recursively organized process between two levels that it uses to combine a cascade of levels. The core process, realized between each of the respective two levels involved in the cascade, consists of one level (input level) that delivers an input for the higher level that functions as a prediction level. The prediction level is supposed to involve some prior expectation, which is fed down to the lower level via feed-back connections and influences the processing on these lower levels. At the relevant lower level, the predicted information and the incoming sensory information are compared, and a possible mismatch is extracted as an error signal. This error signal is communicated via feed-forward to the higher level of processing, where it corrects and adjusts the prior expectation. Via further recurrent loops, a better prior expectation is computed and fed down again to the input level, until a minimal error signal is achieved. In this framework, top-down influences play a crucial role, as they carry the most important information, namely the prediction of how the input information is likely to be composed. This process of adapting the prior expectation until a minimal error signal is reached takes place between all the involved levels of processing.

[9] This argument for cognitive penetration in emotion recognition was first discussed in joint work by Marchi and Newen (2015).

In the case of a perceptual experience, we have prior expectations for each level in a hierarchy of levels. At a rather high level of cognition processing, an estimate percept is connected with the level of expected mental images. Expected mental images can influence top-down the formation of a visual percept. This was recently also shown for cases of binocular rivalry: imagining a specific pattern could strongly bias which of two competing stimuli reach awareness during binocular rivalry (Pearson et al. 2008). Further evidence regarding the neural circuits of top-down processes that supports the framework of predictive coding has been found by Bar (2003, 2009), who proposes a pathway of important top-down modulation that would allow for background knowledge to interfere in the process of forming a percept. He describes a mechanism that, from the perspective of predictive coding, combines a multitude of levels. On his view, the anatomy of the visual system supports quick recruitment of higher-level cognitive areas, such as the orbitofrontal cortex (OFC), before a visual stimulus is recognized (Bar 2003, 2009; Kveraga et al. 2009, 2011). From the OFC, in turn, top-down projections lead to areas in the ventral stream, including recognition areas in the infero-temporal cortex (IT). This illustrates one way in which the top-down modulation of the perceptual experience is implemented. Recently, top-down influences on visual processing have been described in more detail on the neural level (Gilbert and Li 2013) while the neural circuits and their interaction fit the interaction of levels in predictive coding.

In the framework of predictive coding, we can easily account for the claim that prior expectations in the form of mental images can have a top-down influence on the formation of a visual percept. If the expected mental images are relevant patterns of an emotion, then we can nicely re-analyze the experiment of Carroll and Russell (1996) described above. The restaurant story strongly activates a prior expectation of a mental image of a typical face of anger. When looking at the typical Ekman-face of fear, the predictive coding loops start over several levels. A crude description would run as follows: since the prior expectation of an episode of anger is strongly pre-activated by the story; this activates top-down attentional modulations via a cascade of processes, such that those facial input features (of an angry face) that are expected and overlap with the actual face of fear are weighted very heavily, whereas others that are different are weighted very lightly (or inhibited). With such a weighting of facial features, the comparison of expectation and input signal produces only a minor error signal, which is transferred bottom-up. Thus, the expectations need not be adapted. In the process of reiteration, the most probable estimated percept, given the weighted input, is an angry face; thus, the stabilized visual percept (the product of the predictive coding) is the angry face.

If we accept this analysis of the formation of a visual percept, we can conclude that it involves a systematic enrichment of a visual percept by memorized mental images. Thus, it is plausible that our perceptual experiences can have a rich content. But is this only a convincing analysis for the case of emotions, or are we able to generalize to the area of social cognition?

7 Seeing persons by forming person impressions[10]

In our everyday life, the perception of persons and the determination of how we expect them to behave towards us is extremely important. Thus, we rely on the very quick formation of a person impression, which makes it plausible that this is a part of our perceptual experience when seeing persons. To form a person impression (i) we typically pick up some basic features by means of a quick visual evaluation, even when seeing a person for the first time, where (ii) most features are directly associated with socially relevant information, and (iii) they are clustered at the level of perceiving the whole person. Let me offer some support for all three characteristics of the process of forming a person impression, in a situation that is memorized as a person schema (for a detailed description, see Newen 2015).

(i) Quick evaluation, even with parsimonious information: Evaluations of threat (of strong evolutionary relevance) can be made on the basis of exposure to an unfamiliar face lasting as little as 39 ms (Bar et al. 2006). If the exposure to the unfamiliar face lasts for about 100 ms, we are able to evaluate likeability, trustworthiness, competence, and aggressiveness with subjective reliability levels that are similar to those generated under longer viewing times (Willis and Todorov 2006).

(ii) Most features are directly associated with socially relevant information, and this process starts rather early. For example, between 3 and 7 months old, infants learn to distinguish their mother's face from the faces of strangers, and start to categorize people according to emotional expression and sex (Nelson 2001). Here, I am focusing on high-level properties that make the rich content of perceptual experience plausible. Salient biological visual markers allow us to easily register the "big three" categories in person perception (Brewer 1988; Fiske and Neuberg 1990), i.e., sex, race, and age. In the same way, we can illustrate highly informative single features. One important data source here is biological motion-detection, as investigated by point light studies. If a person has lights on her hands, feet, ankles, and some other significant parts of her body, we can videotape her bodily movement in the dark. Such artificial pure biological movement information allows us to register social features, e.g., we can recognize emotions (Ambady and Rosenthal 1992) and attribute personality features (Heberlein et al. 2004) on the basis of seeing dynamic movements alone. There are many more complex culture-dependent visual features that (according to other studies) we use to evaluate the other, e.g., physical attractiveness, where attractive people are evaluated as possessing more desirable characteristics than their less attractive counterparts, a phenomenon that has been labelled the beauty-is-good stereotype (Dion et al. 1972; Eagly et al. 1991). These kinds of stereotypes are especially connected with racial classifications: African-Americans are stereotypically assumed to be lazy, criminal, and uneducated, but also musical and athletic (Devine and Elliot 1995), whereas Asian-Americans are considered to be intelligent, industrious, conservative, and shy (Lin et al. 2005). All these singular features are integrated into person models that enable us to develop detailed and extensive expectations of behavior.

[10] This section draws heavily on Newen (2015).

(iii) Integration of characteristic features at the level of perceiving the whole person: Although I have presented evidence that some single features are very salient for the purpose of transferring social information, there is also considerable evidence suggesting that these features are normally combined with a variety of others to form an integrated impression of a person that I call a person model. Often, gaze alone is not the critical source of information; we actually seem to rely on an integrated evaluation, on the basis of perceiving gaze, head, and body position (Frischen et al. 2007). The same holds for the evaluation of the basic features of sex, race, and age. Although isolated facial features are often sufficient to determine a person's sex, research has indicated that sex categorization is based on the integration of several features (Baudoin and Humphreys 2006; Bruce et al. 1993; Brown and Perrett 1993; Roberts and Bruce 1988; Schyns et al. 2002). Concerning the face, the best available theory of face recognition seems to be Haxby's account (Haxby et al. 2000), according to which there are two distinguishable processes at work, one leading to face identification by focusing more on invariant core features, and the other leading to registering facial expression by relying on varying features. Furthermore, there is evidence that there are two different neural circuits for face perception and body perception (see the review by Macrae and Quadflieg 2010), where both play a core role in registering face or body identity and play an extended role in registering face or body expression in a given situation. And the integration processes are not limited to this level (Martin and Macrae 2007). Since we know that information about facial and bodily features is integrated, e.g., in the evaluation of emotional expression, we can characterize a sequence of integration processes as culminating in a person impression in a situation, which may be stored as a person model in memory. It is argued elsewhere in detail (Newen 2015) that the level of persons is the most important level in integrating social information for human beings.

The process of forming a person impression when seeing a person can be nicely described within the framework of predictive coding. I cannot fully explore this in the present article, but some of the most relevant levels involved in the predictive coding process are highlighted in the overview below (Fig. 4).

8 Conclusion and outlook

The central aim of this article was to defend the liberal-content account for perceptual experiences, with a focus on social perception. I started with non-social perception and argued that the registration of affordances of objects is a perceptual process with a rich content, and that the same holds for expertise perception in the case of chess experts registering the situation when looking at a chess board. Social perception can also result in a percept with rich content, e.g. it can include emotions or person impressions. This was argued while presupposing the metaphysical framework of direct perception: when we perceive the anger of a person, we are related to the anger as a property of a real world entity, not a mental intermediary. It was argued that emotion recognition is a perceptual process and that emotion recognition (even in the case of basic emotions) is rich, since emotions and some other mental phenomena are individuated as integrated patterns of characteristic features, and when perceiving

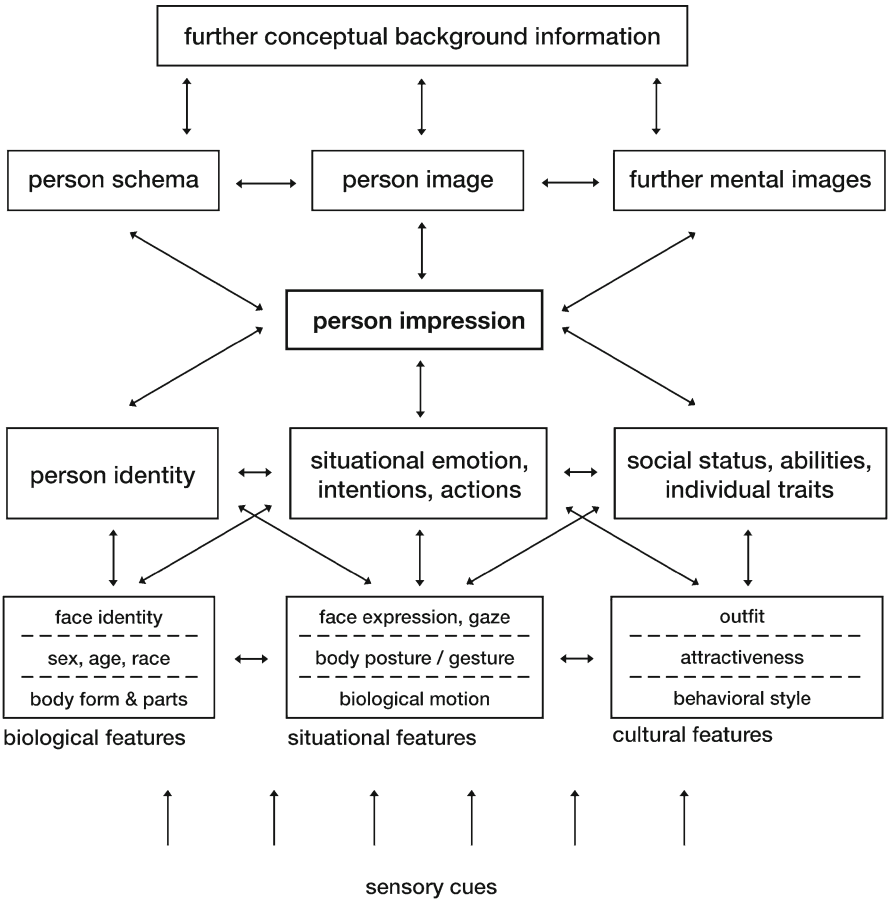

Fig. 4 Figure (from Newen 2015): Forming a person impression is a process of predictive coding leading to a perceptual experience of a person, usually with a rich content: this involves top-down influences from relevant person models of individuals or types of persons (like students, medical doctors, etc.). As developed in detail in Newen (2015), we have to distinguish between implicit and explicit person models, where the former are called person schemata and the latter are called person images

a mental phenomenon the relevant pattern is the content of the experience, often activated by top-down processes as described in predictive coding when the percept is constructed. Thus the content of the perceptual process can be rich, as it is in the case of perceiving basic emotions. Emotion recognition is pattern recognition, where this can be a process of enrichment in relation to parsimonious sensory inputs while remaining a perceptual process; the same holds for forming a person impression when seeing a person. It remains an open question which recognition processes can be described as perception based on pattern recognition processes. As a result of these considerations, we can appreciate that direct perception of mental phenomena (in the sense of having a perceptual experience with the mental phenomenon as its content) is one important strategy in understanding others. But we also rely on other strategies

such as simulation, theory-based inferences, or understanding by interaction. Thus, an important related question for the future of social cognition concerns the exact scope and role of direct perception in social cognition (e.g. Newen 2015). Finally, since the most important mechanism that accounts for enrichment processes is predictive coding, another challenge for social cognition is to work out which phenomena in social cognition rely on this mechanism (e.g. Brown and Brüne 2012; Friston and Frith 2015).

References

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception*, *30*(2), 233–252.

Allen, C. (1999). Animal concepts revisited: The use of self-monitoring as an empirical approach. *Erkenntnis*, *51*(1), 33–40.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 56–274. doi:10.1037/0033-2909.111.2.256.

Avril, M., Leclère, C., Viaux, S., Michelet, S., Achard, C., Missonier, S., et al. (2014). Social signal processing for studying parent–infant interaction. *Frontiers in Psychology*. doi:10.3389/fpsyg.2014.01437.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, *15*, 600–609.

Bar, M. (2009). The proactive brain: Memory for predictions. *Philosophical Transactions of the Royal Society B*, *364*, 1235–1243.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*(2), 269–278. doi:10.1037/1528-3542.6.2.269.

Barlassina, L., & Newen, A. (2013). The role of bodily perception in emotion: In defense of an impure somatic theory. *Philosophy and Phenomenological Research*, *89*(3), 637–678.

Baudoin, J.-Y., & Humphreys, G. W. (2006). Configural information in gender categorisation. *Perception*, *35*(4), 531–540. doi:10.1068/p3403.

Bayne, T. (2009). Perception and the reach of phenomenal content. *Philosophical Quarterly*, *59*(236), 385–404.

Brewer, M. B. (1988). A dual-process model of impression formation. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Mahwah, NJ: Erlbaum.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience*, *6*(147). doi:10.3389/fnhum.2012.00147.

Brown, E. C., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, *22*(7), 829–840. doi:10.1068/p220829.

Bruce, V., Burton, A. M., Hanna, E., Healey, P., Mason, O., Coombes, A., et al. (1993). Sex discrimination: How do we tell the difference between male and female faces? *Perception*, *22*(2), 131–152. doi:10.1068/p220131.

Butterfill, S. A. (2009). Seeing causings and hearing gestures. *Philosophical Quarterly*, *59*(236), 405–428.

Carroll, J. M., & Russell, J. A. (1996). Do facial expression signal specific emotions? Judging emotions from the face in context. *Journal of Personality and Social Psychology*, *70*(2), 205–218.

Chao, L. L., & Martin, A. (2000). Representation of manipulable man-made objects in the dorsal stream. *Neuroimage*, *12*, 478–484.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181–204.

Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton trilogy revisited. *Personality and Social Psychology Bulletin*, *21*(11), 1139–1150. doi:10.1177/01461672952111002.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290. doi:10.1037/h0033731.

Dretske, F. (1981). *Knowledge and the flow of information*. Cambridge, MA: MIT Press.

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but.: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*(1), 109–128. doi:10.1037/0033-2909.110.1.109.

Ekman, P. (1972). *Emotions in the human face*. New York: Pergamon Press.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*, 124–129.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169.

Firestone, C., & Scholl, B. J. (forthcoming). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences*.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology*, *23*, 1–74. doi:10.1016/S0065-2601(08)60317-2.

Fodor, J. (1975). *The language of thought*. New York: Thomas Y. Crowell.

Fodor, J. (2008). *LOT 2: The language of thought revisited*. Oxford: Oxford University Press.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. doi:10.1037/0033-2909.133.4.694.

Friston, K., & Frith, C. (2015). A duet for one. *Consciousness and Cognition*. doi:10.1016/j.concog.2014.12.003.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*(2), 535–543.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, *59*(2), 154–179.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363.

Glock, H. J. (2010). Concepts, abilities and propositions. *Grazer Philosophische Studien*, *81*(2010), 115–136.

Gobet, F. (1997). Roles of pattern recognition and search in expert problem solving. *Thinking and Reasoning*, *3*(4), 291–313.

Gobet, F., & Simon, H. A. (1996). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review*, *3*(2), 159–163.

Griffiths, P. E. (1997). *What emotions really are. The problem of psychological categories*. Chicago: University of Chicago Press.

Griffiths, P. E. (2004). Emotions as natural and normative kinds. *Philosophy of Science*, *71*, 901–911.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Science*, *4*, 223–233. doi:10.1016/S1364-6613(00)01482-0.

Heberlein, A. S., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from pointlight walkers. *Journal of Cognitive Neuroscience*, *16*(7), 1143–1158. doi:10.1162/0898929041920423.

Heelas, P. (1986). Emotion talk across cultures. In R. Harré (Ed.), *The social construction of the emotions* (pp. 234–266). Oxford: Blackwell.

Hohwy, J. (2014). *The predictive mind*. Oxford: Oxford University Press.

Huberle, E., & Karnath, H. O. (2012). The role of temporo-parietal junction (TPJ) in global gestalt perception. *Brain, Structure and Function*, *217*(3), 735–746.

Hurley, S. L. (1998). *Consciousness in action*. Cambridge, MA: Harvard University Press.

Keltner, D., Ekman, P., Gonzaga, G. C., & Beer, J. (2003). Facial expression of emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 415–431). New York: Oxford University Press.

Kveraga, K., Boshyan, J., & Bar, M. (2009). The proactive brain: Using memory-based predictions in visual recognition. In S. Dickinson, M. Tarr, A. Leonardis, & B. Schiele (Eds.), *Object categorization: Computer and human vision perspectives* (pp. 384–400). Cambridge: Cambridge University Press.

Kveraga, K., Ghuman, A. S., Kassam, K. S., Aminoff, E., Hamalainen, M. S., & Chaumon, M. (2011). Early onset of neural synchronization in the contextual associations network. *Proceedings of the National Academy of Sciences*, *108*(8), 3389–3394.

Le Doux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon and Schuster.
Leopold, D., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, *124*(3), 233–251. doi:10.1037/a0019460.
Lin, M. H., Kwan, V. S. Y., Cheung, A., & Fiske, S. T. (2005). Stereotype content model explains prejudice for an envied outgroup: Scale of anti-Asian American stereotypes. *Personality and Social Psychology Bulletin*, *31*(1), 34–47.
Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology*, *3*, 54. doi:10.3389/fpsyg.2012.00054.
Lutz, C. (1986). Emotion words on Ifaluk. In R. Harré (Ed.), *The social construction of the emotions* (pp. 267–288). Oxford: Blackwell.
Macpherson, F. (2012). Cognitive penetration of colour experience. Rethinking the issue in light of an indirect mechanism. *Philosophy and Phenomenological Research*, *84*(1), 24–62.
Macrae, C. N., & Quadflieg, S. (2010). Perceiving people. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology II* (pp. 428–463). Hoboken, NJ: Wiley.
Maley, C. J. (2011). Analog and digital, continuous and discrete. *Philosophical Studies*, *155*(1), 117–131.
Marchi, F., & Newen, A. (2015). Cognitive penetrability and emotion recognition in human facial expressions. *Frontiers in Psychology*, *6*(828). doi:10.3389/fpsyg.2015.00828.
Martin, D., & Macrae, C. N. (2007). A face with a cue: Exploring the inevitability of person categorization. *European Journal of Social Psychology*, *37*(5), 806–816.
Matsumoto, D., & Ekman, P. (1988). Japanese and Caucasian facial expressions of emotion (JACFEE) [Slides]. San Francisco, CA: San Francisco State University, Department of Psychology, Intercultural and Emotion Research Laboratory.
Nanay, B. (2011). Do we see apples as edible? *Pacific Philosophical Quarterly*, *92*, 305–322.
Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*(1–2), 3–18. doi:10.1002/icd.239.
Newen, A. (2015). Understanding others—The person model theory. In T. Metzinger & J. M. Windt (Eds.), *Open MIND 26*. doi:10.15502/9783958570320.
Newen, A., & Bartels, A. (2007). Animal minds and the possession of concepts. *Philosophical Psychology*, *20*(3), 283–308.
Newen, A., & Marchi, F. (2016). Concepts and their organizational structure: Concepts are templates based on mental files. In D. Hommen, C. Kann, & T. Osswald (Hg.), *Concepts and categorization. Systematic and historical perspectives* (pp. 197–227). Münster: Mentis.
Newen, A., Welpinghus, A., & Juckel, G. (2015). Emotion recognition as pattern recognition: The relevance of perception. *Mind & Language*, *30*(2), 187–208.
Noë, A. (2004). *Action in perception*. Cambridge, MA: MIT Press.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 883–917.
Pacherie, E. (2005). Perceiving intentions. In J. Sàágua (Ed.), *A Explicação da Interpretação Humana* (pp. 401–414). Lisbon: Edições Colibri.
Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
Panksepp, J. (2000). Emotion as a natural kind within the brain. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 137–55). New York: Guilford University Press.
Parr, L. A. (2008). Facial expression categorization by chimpanzees using standardized stimuli. *Emotion*, *8*(2), 216–231.
Parr, L. A., & Heintz, M. (2009). Facial expression recognition in rhesus monkeys, *Macaca mulatta*. *Animal Behavior*, *77*(6), 1507–1513.
Pearson, J., Clifford, C., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology*, *18*(13), 982–986.
Proverbio, A. M., Adorni, R., & D'Aniello, G. E. (2011). 250 ms to code for action affordance during observation of manipulable objects. *Neuropsychologia*, *49*, 2711–2717.
Rennig, J., Bilalić, M., Huberle, E., Karnath, H. O., & Himmelbach, M. (2013). The temporo-parietal junction contributes to global gestalt perception-evidence from studies in chess experts. *Frontiers in Human Neuroscience*, *28*(7), 513.
Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception*, *17*(4), 829–840. doi:10.1068/p170475.

Rutherford, M. D., & Kuhlmeier, V. A. (2013). *Social perception: Detection and interpretation of animacy, agency, and intention*. Cambridge, MA: MIT Press.
Schwitzgebel, E. (2002). A phenomenal, dispositional account of belief. *Noûs*, *36*, 249–275.
Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*(5), 402–409. doi:10.1111/1467-9280.00472.
Siegel, S. (2006). Which properties are represented in perception? In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 481–503). Oxford: Oxford University Press.
Siegel, S. (2009). The visual experience of causation. *Philosophical Quarterly*, *59*(236), 519–540.
Siegel, S. (2010). *The contents of visual experience*. Oxford: Oxford University Press.
Siegel, S. (2012). Cognitive penetrability and perceptual justification. *Noûs*, *46*(2), 201–222.
Toribio, J. (2015). Visual experience: Rich but impenetrable. *Synthese*. doi:10.1007/s11229-015-0889-8.
Trujillo, L. T., Allen, J. J. B., Schnyer, D. M., & Peterson, M. A. (2010). Neurophysiological evidence for the influence of past experience on figure-ground perception. *Journal of Vision*. doi:10.1167/10.2.5.
Tye, M. (1995). *Ten problems of consciousness: A representational theory of the phenomenal mind*. Cambridge, MA: MIT Press.
v. Savigny, E. (1988/1989). *Wittgensteins „Philosophische Untersuchungen". Ein Kommentar für Leser*. 2 Bände, Frankfurt a.M.: Klostermann.
van Gelder, T., & Port, R. F. (1995). It's about time: An overview of the dynamical approach to cognition. In T. van Gelder & R. F. Port (Eds.), *Mind as motion: Explorations in the dynamics of cognition* (pp. 1–44). Cambridge, MA: MIT Press.
Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, *27*, 62–75.
Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012a). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, *138*(6), 1172–1217.
Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., et al. (2012b). A century of Gestalt psychology in visual perception: II. *Conceptual and theoretical foundations. Psychological Bulletin*, *138*(6), 1218–1252.
Welpinghus, A. (2015). *Emotions as natural and social kinds*. Münster: Mentis.
Welpinghus, A., & Newen, A. (2012). Emotion und Kultur: Wie individuieren wir Emotionen und welche Rolle spielen kulturelle Faktoren dabei? *Zeitschrift für Philosophische Forschung*, *66*(3), 367–392.
Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598.
Winawer, J., Witthoft, N., Frank, M., Wu, L., Wade, A., & Boroditsky, L. (2007). The Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Science*, *104*(19), 7780–7785.
Zahavi, D. (2011). Empathy and direct social perception: A phenomenological proposal. *Review of Philosophy and Psychology*, *2*(3), 541–558.
Zinck, A., & Newen, A. (2008). Classifying emotions: A developmental account. *Synthese*, *1*, 1–25.