6

Perceptual variation in object perception: A defence of perceptual pluralism

Berit Brogaard and Thomas Alrik Sørensen

1 Introduction

Before reading this chapter, pause for a minute and have a look at your surroundings. Notice the details of the room and the various objects scattered around it. Perhaps you have just brewed a cup of coffee, which now sits next to you on your desk. As you direct your attention towards it, you notice the smell of the coffee and the vapour that slowly ascends from it, conveying to you that it is hot and freshly brewed. Taking a sip, you register the coffee's warmth, liquidity, and smoothness on your tongue and the bitterness of the mocha flavours.

This brief exercise bears witness to our remarkable abilities to rapidly process, categorize, and act on the sensory information we extract from our surroundings. In this chapter, we offer new evidence from philosophy, psychology, and neuroscience that suggests that each of us harbours unique sensory strategies which constitute the basis of our perceptual abilities to categorize sensory input. We devise these strategies to solve the different tasks we encounter in our environment, and our perceptual system structurally adjusts to the strategies that are most successful (Mannix and Sørensen, 2019). Although many of us encounter similar environmental challenges, and therefore devise similar strategies to solve them, strategies can vary substantially across different perceivers. So while you and your colleague may agree that the liquid in the mug is coffee, you may harbour radically different colour, smell, and taste categories for coffee.

On a common characterization, a sensory individual is the output of perceptual categories that represent basic or causally unified features (e.g., Green, 2019). Going by this characterization, our view entails that when different perceivers view the same distal object in the same viewing conditions, their perceptual experiences may represent radically different sensory individuals (i.e., perceptual objects or features).

The thesis we set forth is what we might call 'pluralism about sensory individuals', or 'perceptual pluralism' for short.[1] Perceptual pluralism encompasses the following three

---

[1] As sensory individuals are the products of perceptual categorization, we could also refer to pluralism about perceptual sensory individuals as 'pluralism about perceptual categorization'. We hasten to say that pluralism about sensory individuals should not be confused with pluralism about the format of perceptual content (e.g., iconic vs symbolic) (Quilty-Dunn, 2020).

Berit Brogaard and Thomas Alrik Sørensen, *Perceptual variation in object perception* In: *Sensory Individuals*.
Edited by: Aleksandra Mroczko-Wąsowicz and Rick Grush, Oxford University Press. © Oxford University Press 2023.
DOI: 10.1093/oso/9780198866305.003.0007

propositions: (1) how sensory input is categorized is highly variable across different perceivers as a result of them using different perceptual strategies to solve the perceptual problems they encounter; (2) this variability in perceptual categorization need not be the result of misperception; and (3) contingent on minor constraints (e.g., not being the product of deviant brain processes), the resulting perceiver-specific sensory individuals represent actual features and objects in the external, mind-independent world.

Perceptual pluralism contrasts with what we will call 'perceptual monism'. To a first approximation, this is the view that for any sensory input, there is a single correct way to categorize it. We take this view—in its most extreme iteration—to imply that correctly processed and categorized sensory individuals represent external, mind-independent features or objects in a 1:1 fashion.

Before proceeding, we should clarify what we mean by 'perceptual strategies'. Perceptual strategies should be kept apart from the cognitive tasks they subserve. For example, an important cognitive task in everyday life is that of recognizing other people. To solve this task, one may apply a range of different perceptual strategies, for example, the recognition of facial features, gait, clothing, voice, context, etc. (Sørensen and Overgaard, 2018). With increased expertise, some strategies may be more successful than others (e.g., using facial features for identification). In a Hebbian fashion (Hebb, 1949), the brain may increasingly rely on these strategies. The array of strategies of a particular perceiver thus depends on her specific expertise.

Acquiring expertise involves a learning process, typically in the form of repetition and practice, rather than mere instruction. Some forms of expertise require enhanced abilities to remember, imagine, or control the body. For example, excelling at memory-matching smartphone games requires developing enhanced memory skills; excelling at mental rotation games requires developing enhanced imaginative skills, and excelling at line dancing requires enhanced bodily coordination and balancing skills. A different kind of expertise is perceptual expertise. Learning processes that result in perceptual expertise—that is, perceptual learning—enhance perceptual capacities by altering the structure or function of the perceptual system. Three distinct perceptual learning mechanisms have been identified (Chudnoff, 2018; Goldstone, 1998), which we will call 'differentiation', 'unitization', and 'anchoring'.[2]

2 Differentiation as a perceptual learning strategy

Differentiation alters the attentional weight assigned to parts or features of a stimulus (Cecchi, 2014; Ransom, 2020). On the level of the brain, the receptive fields of the neurons that initially encode the stimulus adapt to become more sensitive to important features and structural relations (Gilbert and Sigman, 2007).

This kind of learning was first examined in Gibson and Gibson's (1955) classic pattern-matching experiment. The participants (adults, children aged 8.5–11, and children aged 6–8) were presented with a target card depicting a nonsense scribble for 5 seconds and then a deck of 34 cards, one at a time, for 3 seconds, to determine for each whether it was a match.

---

[2] An older categorization divides perceptual learning into: 'imprinting', 'attentional weighting' (differentiation), and 'chunking' (unitization) (see, for example, Goldstone (1998)). According to Goldstone (1998), imprinting is the generation of an internal model, or memory representation, based on repeated exposure to a stimulus.

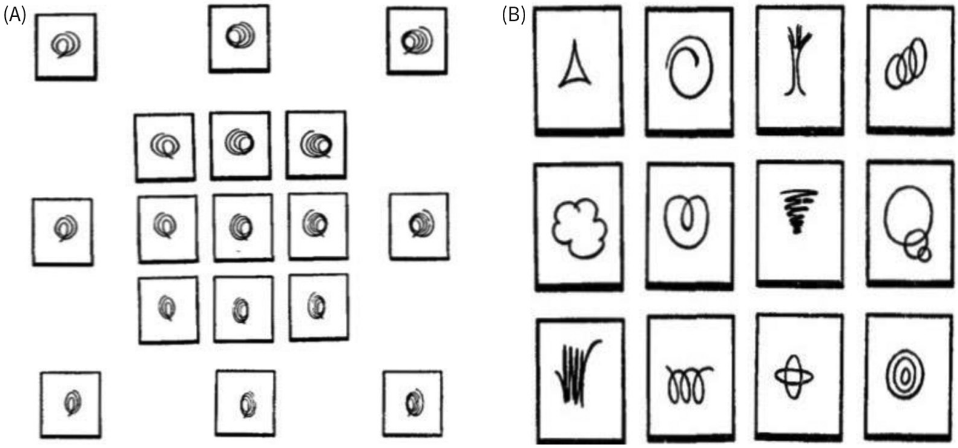

Figure 6.1 (A) Nonsense scribbles differing in three dimensions of variation (i.e., number of coils, compression, and orientation). (B) Nonsense scribbles differing in many dimensions of variation.
From Gibson and Gibson, 1955.

Seventeen of the 34 cards were highly similar to the target (Figure 6.1A); 12 were easily distinguishable from the target (Figure 6.1B), and four matched the target exactly. This process was repeated (without feedback on accuracy) until the participants made the four correct identifications in a single trial.

As the results show, each age group had a different learning curve (Table 6.1). For the adult group, the number of undifferentiated items at the outset was small (mean = 3.0), and they only needed a few trials to identify the target (mean = 3.1). For the youngest group, the number of trials needed to identify the target could not be fully determined. The older

Table 6.1 Performance across three age groups

Variable	**Adults** ($N = 12$)	**Older children** ($N = 10$)	**Younger children** ($N = 10$)
Mean number of undifferentiated items on first trial	3.0	7.9	13.4
Mean number of trials required for completely specific response	3.1	4.7	6.7*
Percentage of erroneous recognitions for items differing in *one* quality	17	27	53
Percentage of erroneous recognitions for items differing in *two* qualities	2	7	35
Percentage of erroneous recognitions for items differing in *three* qualities	0.7	2	28

* Only two of the younger children achieved a completely specific identification. The mean number of undifferentiated items on the last trial ($n = 6.7$) was still 3.9.

From Gibson and Gibson, 1955.

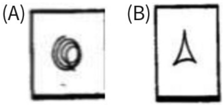

Figure 6.2 A target scribble (A) and random scribble (A). (Adapted from Figure 6.1A.)

children performed somewhere in between, with a mean of 7.9 undifferentiated items on the first trial and a mean of 4.7 trials required to identify a target.

Prior to the learning process, participants were able to distinguish the centre scribbles (Figure 6.2A) from a clear outlier (Figure 6.2B). But after the learning process, the older children and adults had acquired the ability to distinguish between pairs of scribbles, which only differed in the number of coils, compression, and orientation (Figure 6.2A).

The experiment thus demonstrates that detected (or detectable) features of a stimulus (e.g., number of coils, compression, orientation) can come to serve as a diagnostic for recognition after a perceptual learning process.

3 Unitization as a perceptual learning strategy

Unitization, the second variety of perceptual learning, can be viewed as the opposite of differentiation. Whereas differentiation divides wholes into more differentiated features, unitization integrates parts into units. An example of expertise that depends on unitization can be found in chess experts. Studies comparing experts and novices indicate that while novices can only encode the position of individual chess pieces in short-term memory, experts are able to encode whole chess configurations (e.g., Gobet and Simon, 1996). The basic unit encoded is the 'chunk' (Miller, 1956; Wenger and Rhoten, 2020), which consists of a configuration of pieces that are frequently encountered together and that are related by type, colour, role, and position (Chase and Simon, 1973a, b).

There seem to be no patent differences in thought process statistics (e.g., number of moves considered, search heuristics, etc.) between expert and novice chess players (de Groot, 1966). However, studies have found differences in perceptual processing between experts and novices: experts are able to recall a chess position almost perfectly after viewing it for 5 seconds. What seems to account for this difference is not the experts' superior memory, but their ability to perceive the structure of certain positions and encode them as units. Similar expertise-related enhancements of short-term memory have also been reported with more general stimuli (e.g., Sørensen and Kyllingsbæk, 2012).

Chase and Simon (1973b) tested the hypothesis that expert chess players extract more information from a chess position than less skilled players, using two tasks. In the perceptual task, players were asked to reconstruct a position in plain view as quickly and as accurately as possible. The players were assumed to encode one chunk per glance. The number and duration of glances at the board were used as an indicator of the size of chunks encoded in memory. In the memory task, players were asked to reconstruct a position from memory following a 5-second exposure. If they did not reconstruct the position correctly on the first try, their board was cleared and the task repeated with the same position until they got it right. The time between the placement of successive pieces was used as an indicator of the size of the chunks. The findings indicated that the experts performed superiorly because of their ability to encode positions into larger perceptual chunks.

In a separate experiment, Chase and Simon (1973a) asked beginners, Class A players, and Masters to first memorize the steps in a game with 25 positions and then to reconstruct a specific game position from memory. The initial hypothesis was that novice players would rely more on the function of individual pieces (e.g., an attack) than spatial patterns of positions. This was confirmed by the results, which showed that beginners had limited access to larger patterns in long-term memory: 80% of novices' retrieved patterns consisted of a single piece, indicating that they had to reconstruct positions one piece at a time. Masters and Class A players recalled patterns consisting of only a single piece, only 26% and 37% of times, respectively, suggesting access to larger patterns in long-term memory. The more skilled players were furthermore able to retrieve encoded configurations up to four times faster than beginners.

A study by Reingold *et al.* (2001) points to the enhanced skill set of expert chess players being perceptual rather than conceptual. Here, a minimized 5 × 5 chessboard was displayed to novice, intermediate, and expert chess players. In session one, configurations fell into two types: (1) configurations with 2–3 pieces in a checking ('yes') condition (e.g., the bishop in one corner, the king in the diagonal corner); (2) configurations with 2–3 pieces in a non-checking ('no') condition (e.g., the rook in one corner, the king in the diagonal corner) (Figure 6.3).

The players were told to determine as quickly and as accurately as possible whether the king was in check. The results showed that non-experts responded more slowly when there

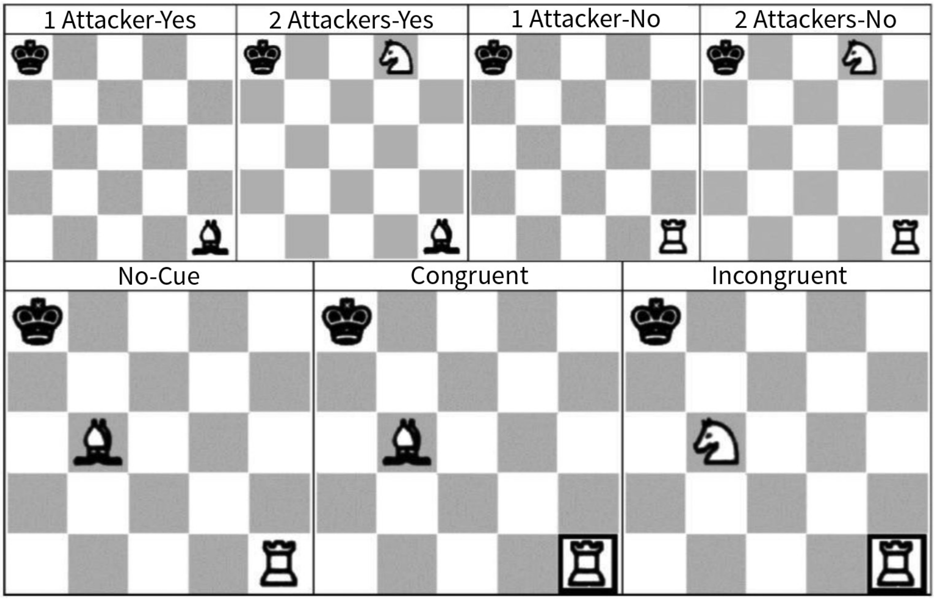

Figure 6.3 Examples of the check configurations. The top row demonstrates 'yes' (check) vs 'no' (non-check) conditions. The bottom row illustrates the no-cue condition ('no' trials) and conditions in which a cued non-checking attacker appeared together with an attacker that was either congruent (i.e., non-checking) or incongruent (i.e., checking).
From Reingold *et al.*, 2001.

were two attackers than one. The extra piece did not affect experts, which indicates holistic processing for experts, but non-holistic processing for non-experts. The finding points to the expert chess players possessing enhanced perceptual skills, enabling them to process chess configurations perceptually as units.

In the second session, only the two-attacker positions (e.g., bishop/rook plus the king) from the first part were used and double-check positions were added to create four possible combinations of checking for both attackers. The non-checking configuration was the congruent condition, and the checking configuration the incongruent condition. On half of the trials, one of the attackers was cued (in red). Participants were instructed to proceed as before if there was no cue. If a cue was present, they should ignore the non-coloured attacker. If the configurations were processed as individual piece, cueing should improve performance because the player would not need to examine the non-cued checking configuration. But if the configurations were processed holistically, cueing should not improve performance. The results showed that cueing only helped experts, suggesting that they process chess configurations holistically. Surprisingly, when a cue was present, experts were faster in the congruent (non-checking) than the incongruent (checking) condition. This may be due to a Stroop-like interference (Stroop, 1935), owing to the incongruent (check) relation that was to be ignored grabbing the experts' attention. In expert chess players, the Stroop-like interferences may occur because experts' enhanced skills, which enabled them to perceive chess configurations holistically, interfered with tasks requiring non-holistic processing of individual chess pieces. Thus, training to become an expert alters how we perceptually process stimuli.

4 Anchoring as a perceptual learning strategy

Anchoring, the third type of perceptual learning, improves performance on perceptual categorization (or recognition) tasks by linking a target category and an unrelated, but biologically more fundamental or personally more significant, category (Hubel and Wiesel, 1962). Anchoring can serve as a mnemonic for learning meaningless information (e.g., random strings of numbers) (e.g., Elmera *et al.*, 2017; Sutherland, 2017). As we will see, it also lies at the core of the perceptual abnormality known as 'synaesthesia'.

Synaesthesia is an unusual way of perceiving the world in the healthy population that involves automatic connections between seemingly unrelated experiences, images, or concepts. For example, the number 3 printed in black ink may look like it is covered with a thin film of copper green, the word 'street' may flood the mouth with the flavour of lentil soup, and the key of C# minor may elicit a bright pink spiral radiating from the centre of the visual field. In grapheme–colour synaesthesia, one of the most common forms, perceiving or thinking about an achromatic grapheme (known as the 'inducer') gives rise to the experience (known as the 'concurrent') that the grapheme has a specific colour (Armel and Ramachandran, 1999; Ásgierson, Nordfang, and Sørensen, 2015; Brogaard, 2014; 2020; Grossenbacher and Lovelace, 2001).

Most cases of synaesthesia are developmental, starting in early childhood. Whether synaesthesia is a hereditary trait has been the subject of fierce debate, partly because the question of its genetic basis has been thought to be intertwined with that of whether the condition can be learned. Evidence cited in favour of the heredity of synaesthesia includes family clustering data (Baron-Cohen *et al.*, 1996) and family linkage studies (e.g., Tilot *et al.*, 2018).

As family members often are exposed to similar environmental factors, family clustering data fail to demonstrate heredity (Brogaard, Marlow, and Rice, 2014; 2016; Mannix and Sørensen, 2019; Sørensen, 2019). Family linkage studies provide more compelling evidence for a genetic basis in some forms of synaesthesia (e.g., Tomson, Avidan, Lee, *et al.*, 2011). Genetic markers from family linkage studies have also been implicated in developmental pruning errors (Carmichael and Simner, 2013), suggesting that synaesthesia may sometimes occur when insufficient pruning results in atypical information processing.

But even if all forms of synaesthesia should turn out to involve a hereditary component, this would not preclude that learning plays a central role in its genesis. This is because even if a condition is hereditary, it is not inevitable that a genetically predisposed person will develop it. For example, a variety of personal and environmental factors have been found to contribute to the development of breast cancer, including, gender, age, ethnicity, diet, lifestyle, and hormonal and reproductive factors. Likewise, several personal and environmental factors have been found to contribute to the development of synaesthesia, including normal tendencies to associate certain items, for instance, high pitch and lighter colours (Mondloch and Maurer, 2004), and pitch and taste (Holt-Hansen, 1968), as well as synaesthesia-specific effects like the inducer's visual or spectral shape (Brang and Ramachandran, 2011), the inducer's contextual frequency (Beeli *et al.*, 2007; Simner *et al.*, 2005), and reliance on inducer–concurrent connections during early learning (e.g., Witthoft, Winawer, and Eagleman, 2015).

Witthoft and Winawer (2013) matched data from 11 grapheme–colour synaesthetes who had completed a synaesthesia battery. Ten of the 11 subjects reported having owned one of three sets of refrigerator magnets with the same colour scheme sold by FisherPrice between 1972 and 1989. The researchers found a remarkable similarity of letter–colour associations among the synaesthetes. The fewest matches among the synaesthetes were 14 letters. The probability of finding 14 or more letter matches in 26 chances is less than one in a billion, which strongly suggests that the subjects acquired their letter–colours associations from the Fisher-Price toys. The toys may thus have played a role in the development of grapheme–colour synaesthesia because the formation of synaesthetic connections between the letters and colours turned out to be a successful strategy for learning to categorize letters (Mannix and Sørensen, 2021). Being a biologically more fundamental category than letters, the colours may have served as anchors for the categorization of letters.

This raises the question of whether neurotypical adults can be trained to become synaesthetes. Studies have shown that when non-synaesthetes are repeatedly exposed to reading materials with letters printed in different colours, and subsequently are asked to name the grapheme ink colours that are incongruent with their synthetic colours, this results in Stroop interference, a mark of grapheme–colour synaesthesia (e.g., Bor *et al.*, 2014).

However, these studies did not induce genuine grapheme–colour synaesthesia. First, the incongruency effects were typically smaller than those found in grapheme–colour synaesthesia (Elias *et al.*, 2003). Second, neuroimaging studies have shown that trainees do not have the same brain activation patterns as synaesthetes (Nunn *et al.*, 2002). The limited success of training studies could be due to the lack of a genetic basis for synaesthesia in participants. Alternatively, it may have been that learning to associate graphemes with colours failed to serve as a strategy for mastering a useful or personally significant task, such as acquiring the alphabet, which is only a useful task for school-aged children (Brogaard, 2017; 2020; Mannix and Sørensen, 2021). Finally, although several studies provide extensive training, this is still not on a par with the amount of training required for learning new skills (e.g., reading).

Sørensen and Kyllingsbæk (2012) raised a similar concern about studies claiming that training does not affect short-term memory capacity. These studies apparently demonstrated that extensive training did not affect short-term memory capacity (e.g., Chen *et al*., 2006). However, in Sørensen and Kyllingsbæk (2012), we used a change–detection paradigm (Luck and Vogel, 1997) to examine short-term memory capacity for letters and line drawings across age groups—the logic being that the different age groups had increasing familiarity and expertise with letters, whereas none of the participants had any overt training with the line drawings. Across the different groups, there was no change in memory capacity for line drawings. The youngest age group was only able to retain two letters, but with increasing age, we saw a systematic capacity increase up to adulthood where participants could retain approximately four letters. This pattern was replicated in a slightly different design by Dall *et al*. (2016) in adults, who demonstrated stable short-term memory capacity for both line drawings and letters. Nevertheless, in a critical condition using Japanese hiragana, there was a systematic short-term memory increase across subjects who did not know Japanese, studied Japanese, or were experts, corroborating the link between expertise and short-term memory capacity.

As we saw, synaesthesia can be learnt from environmental cues, which may reflect prioritization of a particular perceptual strategy that can be applied during learning. The notion that we develop different perceptual strategies to enhance learning is supported by the fact that grapheme–colour synaesthesia is one of the most common types (Ward, 2013). It is not far-fetched to imagine that when faced with the task of learning new abstract categories (e.g., letters), linking the new categories to a previously established category (i.e., a colour) would aid learning (Mannix and Sørensen, 2021).

Evidence for this theory also comes from the observation that synaesthetic colours sometimes change (e.g., Sørensen, Nordfang, and Ásgeirsson, 2016), emerge over time (Simner and Bain, 2013), or—anecdotally at least—completely disappear during the teen years or in early adulthood (Keane, 2017). As synaesthetes would have learnt to read, write, and manipulate numbers at this age, the synaesthetic associations would have lost their usefulness as learning strategies. Relatedly, the number '0' often lacks a colour or has an unusual colour (e.g., gold) in young synaesthetes until they reach school age (Brogaard, Marlow, and Rice, 2014), reflecting that most children do not develop a good understanding of the meaning of '0' until they face maths problems in school. At this point, strategically associating '0' with a basic hue may facilitate grasp.

5 Perceptual categorization and expertise

Perceptual processing goes through various stages during which sensory information is interpreted and encoded into memory. Goldstone (1998) regards learning based on the imprinting of sensory information in memory and the repurposing of this information in categorization as a kind of perceptual learning. Traditionally, perceptual processing has been described as an initial stage where attention prioritizes sensory information for encoding in short-term memory, followed by a stage of extended processing which increases the probability of encoding into more durable long-term memory representations (e.g., Atkinson and Shiffrin, 1968). Although this model was criticized by clinical studies in the early 1970s, it aligns with a folk psychological interpretation of how information flows from short-term memory to long-term memory.

An alternative to the traditional model holds that sensory information initially is matched with object and feature representations in long-term memory (Brogaard and Sørensen, in press), and that the most relevant representations then are selected for representation in short-term memory. The theory of visual attention (TVA; Bundesen, 1990) proposes this type of relationship. According to the TVA, the sensory input is modulated by two attentional top-down parameters prioritizing selection for features (filtering) or objects (pigeonholing). This model allows for flexible parallel attentional prioritization based on filtering and pigeonholing, as well as the match between the object/features and the perceiver's memory templates (for a formal mathematical description, see Bundesen, 1990). In fact, the mechanisms for perceptual learning we described earlier bear resemblances to the TVA's three core principles of selection: feature selection (*differentiation*), object selection (*unitization*), and template matching between sensory input and long-term memory (*anchoring*).

If models like the TVA that assume a more central role for long-term memory in perceptual processing are correct, then we should expect the consolidation of new perceptual categories to facilitate the matching of sensory signals to these categories. Differences in perceptual processing as new categories are acquired in a variety of different domains seem to support this notion, for example, face processing (Curby and Gauthier, 2007), car experts (Curby et al., 2009), and cartoons (e.g., Xie and Zhang, 2017).

Recently, we investigated two aspects of complexity: one varying number of feature elements and another varying stimulus familiarity (Dall et al., 2021). Jackson and Raymond (2005) described this distinction as one between physical (feature elements) and perceived complexity (familiarity). Using the TVA to isolate the specific components of attention, we showed that Chinese expert readers[3] are not affected by early parameters like threshold for visual perception. We nevertheless saw a systematic influence of familiarity on processing speech and short-term memory capacity. Interestingly, this modulation of processing speed and memory depended only on familiarity and not at all on the increase in feature elements. So this latter study points to a more holistic processing of Chinese characters by experts, akin to the holistic processing of other stimulus types (e.g., faces or chess configurations). For novices, the number of feature elements seemed to be the driving factor.

In light of the previously discussed findings supporting holistic processing of chess configurations in experts, this evidence indicates that memory representations influence how sensory information is perceptually categorized. These memory representations, in turn, depend on prior perceptual encounters with the relevant categorization tasks.

Like prior experience, the languages we speak can also shape perceptual categorization. According to the category adjustment model of colour memory, colour representations in short-term memory are the result of combining information in different proportions from two sources (Cibelli et al., 2016). One is the fine-grained bottom-up representation of the colour seen, and the other is a coarse-grained representation of colour as depicted by a language-specific category (e.g., blue). The combination of the two sources of information results in a colour representation in short-term memory that is modulated by a language-specific colour category.

---

[3] Combining a whole report design (Sperling, 1960) with varied exposures, Dall *et al.* (2021) were able to fit the TVA model to the exponential performance of Chinese experts and extract a range of attentional parameters, thus investigating Jackson and Raymond's (2005) processing dimensions. Stroke count in Chinese characters were used to capture physical complexity and character frequency as a measure of perceived complexity over four variables: low vs high stroke count and low vs high familiarity.

The TVA offers a plausible explanation of the category adjustment model. If, as the TVA suggests, incoming colour signals are initially matched with colour categories in long-term memory, and the most relevant colour categories are selected for representation in short-term memory, then we should expect the fine-grained bottom-up colour representation to be modulated by language-specific colour categories.

This, in turn, points to a form of linguistic relativity. The Sapir–Whorf hypothesis, a strong form of linguistic relativity, states that the languages we speak can modulate our perceptual, affective, and cognitive states. While we may not fully align with the strong interpretation of this hypothesis, the TVA and the category adjustment model suggest that prior knowledge modulates perceptual categorization. In the Section 6, we explore how colour language can influence early perceptual processing.

6 Linguistic effects on colour categorization

Originally set forth by Sapir (1921) and Whorf (1956), the hypothesis that our native language shapes what we think, remember, and perceive is also known as 'linguistic relativity'. Linguistic relativity about colour categorization came under harsh criticism in the late 1960s when Berlin and Kay (1969) popularized the view that despite documented cultural variation in colour language (e.g., Bornstein, 2007), there is a set number of basic colour categories. Writing during an era when academics increasingly believed that language and cognition were universal, innate phenomena,[4] Berlin and Kay maintained that their basic colour categories were innate, and hence culturally ubiquitous. Cultural variation in colour language, they argue, reflect that cultural processes influence which colour terms people use to depict the basic colour categories.

Although few thinkers are in the grip of Berlin and Kay's universality thesis today, the Sapir–Whorf hypothesis remains highly controversial (e.g., Regier and Xu, 2017). This is so despite relatively sparse evidence against it. Historically, the main challenge to the Sapir–Whorf hypothesis has been the finding that speakers of languages with very few colour words sometimes perform as well as English speakers on colour categorization tasks, which would suggest that non-English speakers' perceptual colour categories align with those of English speakers (Rosch, 1978).

However, studies in psycholinguistics have since then provided compelling evidence for thinking that the linguistic concepts people possess in their native language can modulate early perceptual processing (Athanasopoulos et al., 2015; Thierry et al., 2009). Thierry et al. (2009) demonstrated this effect in native English and Greek speakers. Greek contains two unique hue concepts for blue: '*ghalazio*' and '*ble*', corresponding to what native English speakers regard as two shades of the unique hue blue (Figure 6.4).

Using electroencephalography, the team recorded brain waves from the skull of the participants while they were performing a basic oddball shape discrimination task in which they had to press a button when seeing a square shape (the target) within a regularly paced stream of circles (Figure 6.5).

---

[4] Universality theses that gained enormous influence during the 1960s include most famously Chomsky's generative grammar, but also Ekman's six basic emotional categories.

Figure 6.4 Greek contains two colour concepts for blue: darker blue *ble* (left) and lighter blue *ghalazio* (middle).

This discrimination task helped keep the participants' attention focused on something other than what was measured, namely the strength of the brain's unconscious response to a colour change during a trial. In one block of trials, circles were *ghalazio* by default, and 'oddballs' were *ble*. In the second block, this pattern was reversed, and in the final two 'control' blocks, the experiments from the first two blocks were repeated, but with light and dark green figures.

The findings showed a significantly greater brain wave response to a colour shift from *ghalazio* to *ble*, or vice versa, in Greek vs English speakers. A stronger response to a colour shift signifies a greater perceptual difference between the colours.

Importantly, responses to a change from light green to dark green, or vice versa, were found to be the same in English and Greek speakers. Moreover, the Greek participants' responses to a shift in a shade of green were significantly smaller than their responses to a shift in a shade of blue. These findings provide strong support for the claim that not only does possessing the concepts *ghalazio* and *ble* shape how Greek speakers memorize and think about the two colour shades, but it also modulates colour categorization at the lowest level of perceptual processing.

These findings suggest that very early perceptual processing is shaped by linguistic concepts we acquired as children. This, in turn, provides renewed credence to the Sapir–Whorf hypothesis. Taken together with the previous sections, this points to a highly individualized

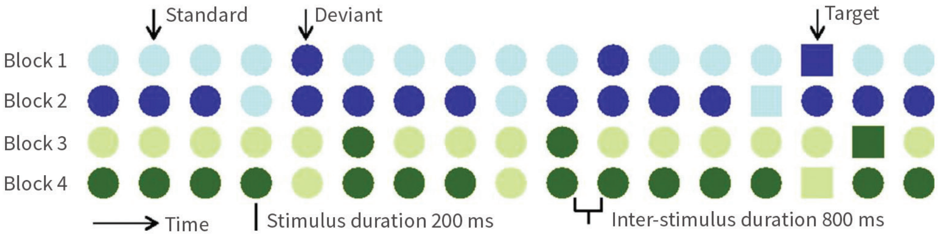

Figure 6.5 Participants were asked to press a button when they saw a square within a regularly paced stream of circles. Meanwhile, the researchers measured the visual mismatch negativity effect in response to a light colour/dark colour transition.
From Thierry *et al.*, 2009.

variability in perceptual categorization in different domains, including, but not limited to, language.

7 Perceptual pluralism defended

The evidence reviewed above jointly indicates that how sensory information is processed and categorized can vary greatly across perceivers relying on different perceptual strategies to solve the perceptual tasks they encounter. This is the first conjunct of perceptual pluralism, which—as noted in the introduction—we take to encompass the following three claims: (1) how sensory input is categorized is highly variable across different perceivers as a result of them using different perceptual strategies to solve the perceptual problems they encounter; (2) this variability in perceptual categorization need not be the result of misperception; and (3) contingent on minor constraints (e.g., not being the product of deviant brain processes), the resulting perceiver-specific sensory individuals represent actual features and objects in the external, mind-independent world.

As perceptual pluralism encompasses all three claims, it should come as no surprise that the first claim (1) is perfectly compatible with perceptual monism, the view that there is only one correct way of categorizing perceptual information. If indeed there were a single correct way to categorize sensory input, then the empirical evidence reviewed above would merely be documenting that we are accurately categorizing sensory input in some circumstances, but not in others, which would be old hat.

Several philosophers have made a case for some restricted version of perceptual monism (although not under that name). For example, Byrne (2006) offers a defence of what we term perceptual monism which is restricted to colour perception. According to Byrne, there are objective facts about the colour of objects, but there is irresolvable disagreement about what those facts are. Byrne (2006, p. 337) expresses the view as follows (in response to Cohen, 2006):

> Suppose that [two] normal human observers Mya and Bo are viewing a chip C . . . C looks unique green to Mya, and bluish green to Bo. The problem, as Cohen has it, is to explain 'what would (metaphysically) make it the case' that Mya, say and not Bo, is perceiving C correctly. [Cohen] purports to find the explanation 'extremely hard to imagine', and so concludes that both Mya and Bo are perceiving C correctly. . . . [I maintain that] what 'makes it the case' that Mya, not Bo, is perceiving C correctly, is that Mya is representing C as being unique green, Bo is representing C as being bluish green (no problem so far), and C is unique green, not bluish green (likewise no problem).[5]

Byrne's view is this: whenever two perceivers disagree about the colour of an external, mind-independent object or about whether two external, mind-independent objects have the same colour, at least one of them is wrong. One apparent problem with this view is that it entails that there are unknowable colour facts. For any coloured object, there are bound to be individuals with no apparent visual system defects who would disagree about its colour, were they to view it in the same viewing conditions. Yet we would have no basis for deciphering

---

[5] We have replaced S_1 with Mya, and S_2 with Bo.

who is right and who is wrong. So for any coloured object, the truth about its colour would be unknowable. This form of radical colour epistemicism seems highly implausible.

Our objection to Byrne generalizes to similar defences of other types of perceptual monism: generally speaking, it is implausible to think that, for any sensory input, people fall neatly into groups of normal and abnormal perceivers, and it is also improbable that perception gives us no reliable insight into the structure of the external, mind-independent world.

Similarly, our objections lend credence to the second claim of perceptual pluralism, viz. that perceptual variability in a population of speakers regarding the same sensory input need not be the result of misperception, which is to say that overlapping and disjoint categorizations of a given perceptual input may all represent real categories in the external, mind-independent world (see also Cohen, 2009).

Thus, the perceptual categorization of a certain sensory input as the unique hue blue in native English speakers represents an actual category of blue objects. Likewise, the perceptual categorization of the very same sensory input as the two unique hues *ghalazio* and *ble* in native Greek speakers represents two real categories of coloured objects.

The same can be said for other variations in perceptual categorization. In Gibson and Gibson's (1955) pattern-matching experiment, the untrained participants categorized the scribbles as belonging to a single category of 'scribbles' (Figure 6.1A). But after training, the same participants made more fine-grained categorizations of the scribbles. Likewise, in the studies of chess expertise, experts and novices were found to categorize sensory input from chess pieces on a chessboard in different ways (e.g., Chase and Simon, 1973a). Novices classified the sensory input in terms of individual chess pieces, whereas experts classified the same input in terms of holistic chess configurations. But while we can say that the chess experts are better at chess, we cannot conclude that only the experts' perceptual categorizations are the correct ones—they are merely different due to the modulation of expertise.

Finally, we argued that anchoring is the predominant perceptual learning mechanism in synaesthesia. Synaesthetes improve on categorization tasks by creating perceptual connections or semantic memory associations between a target category and an unrelated category that is biologically more fundamental or of greater personal significance. Consider, for example, a grapheme–colour synaesthete who connects single digits with unique colours. For her, the number '3', say, does not just belong to a mathematical category, but also to the category of green objects.

8 Conclusion

Here we have defended pluralism about perceptual categorization based on studies suggesting that variability in perceptual categorization and the resulting sensory individuals reflects that we use different strategies for solving the different tasks we encounter in our environment, and that the perceptual system functionally and structurally adjusts to the strategies that are most successful.

Perceptual pluralism entails that prior sensory input encoded in memory modulates our perceptual objects (or sensory individuals). This may raise a concern about whether perceptual pluralism entails that low-level perception is cognitively penetrable (e.g., Firestone and Scholl, 2016; Pylyshyn, 1999). Controversies about this thesis concern its methodological basis (Brogaard and Gatzia, 2017; Firestone and Scholl, 2016), its implications for a popular view about perceptual justification (Chudnoff, 2020; Siegel, 2017), and its implications in perceptual bias (Brogaard, 2021; Siegel, 2017).

Rather than providing a definitive answer to the question of whether perceptual pluralism entails the cognitive penetrability thesis, however, our argument shows that not much hinges on how we answer it, because whatever disquieting conclusions can be derived from it can also be deduced, given more innocuous top-down influences.

Acknowledgements

We are grateful to two reviewers for invaluable comments on a previous version of this paper. Sørensen was supported by the Independent Research Fund, Denmark. Brogaard was supported by a Cooper Fellowship from the University of Miami.

References

Armel, K. C. and Ramachandran, V. S. (1999). Acquired synesthesia in Retinitis Pigmentosa. *Neurocase*, 5, 293–96.

Ásgeirsson, Á. G., Nordfang, M., and Sørensen, T. A. (2015). Components of attention in grapheme–color synesthesia: a modeling approach. *PLoS One*, 10(8), e0134456.

Athanasopoulos, P., Bylund, E., Montero-Melis, G., *et al*. (2015). Two languages, two minds: flexible cognitive processing driven by language of operation. *Psychological Science*, 26(4), 518–26.

Atkinson, R. C. and Shiffrin, R. M. (1968). Human memory: a proposed system and its control processes. In *The Psychology of Learning and Motivation: Advances in Research and Theory*, edited by K. W. Spence and J. T. Spence, 89–195. New York: Academic Press.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., and Bolton, P. (1996). Synaesthesia: prevalence and familiality. *Perception*, 25(9), 1073–9.

Beeli, G., Esslen, M., and Jäncke, L. (2007). Frequency correlates in grapheme–color synaesthesia. *Psychological Science*, 18(9), 788–92.

Berlin, B. and Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press.

Bor, D., Rothen, N., Schwartzman, D. J., Clayton, S., and Seth, A. K. (2014). Adults can be trained to acquire synesthetic experiences. *Scientific Reports*, 4, 7089.

Bornstein, M. H. (2007). Hue categorization and color naming: cognition to language to culture. In *Anthropology of Color: Interdisciplinary Multilevel Modeling*, edited by R. E. MacLaury, G. V. Paramei, and D. Dedrick, 3–27. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Brang, D. and Ramachandran, V. S. (2011). Survival of the synesthesia gene: why do people hear colors and taste words? *PLoS Biology*, 9(11), e1001205.

Brogaard, B. (2014). Varieties of synesthetic experience. In *Phenomenology and the Neurophilosophy of Consciousness*, edited by R. Brown, 409–12. Springer.

Brogaard, B. (2017). Synesthetic binding and the reactivation model of memory. In *Sensory Blending: On Synaesthesia and Related Phenomena*, edited by O. Deroy, 126–50. Oxford: Oxford University Press.

Brogaard, B. (2020). Color synaesthesia and its philosophical implications. In *The Routledge Handbook of Philosophy of Colour*, edited by D. Brown and F. Macpherson. Francis & Taylor.

Brogaard, B. (2021). Implicit biases in visually guided action. *Synthese*, 198, 3943–67.

Brogaard, B. and Gatzia, D. E. (2017). Is color experience cognitively penetrable? *Topics in Cognitive Science*, 9(1), 193–214.

Brogaard, B., Marlow, K., and Rice, K. (2014). The long-term potentiation model for grapheme–color binding in synesthesia. In *Sensory Integration and the Unity of Consciousness*, edited by D. Bennett and C. Hill, 37–72. Cambridge, MA: MIT Press.

Brogaard, B., Marlow, K. and Rice, K. (2016). Do synesthetic colors grab attention in visual search? *Review of Philosophy and Psychology*, 7(4), 701–14.

Brogaard, B. and Sørensen, T. A. (in press). Predictive processing and object recognition. In *Expected Experiences: The Predictive Mind in an Uncertain World*, edited by T. Cheng and J. Hohwy. New York, NY: Routledge.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97(4), 523–47.

Byrne, A. (2006). Comments on Cohen, Mizrahi, Maund, Levine. *Dialectica*, 60, 337–40.

Carmichael, D. A. and Simner, J. (2013). The immune hypothesis of synesthesia. *Frontiers in Human Neuroscience*, 7, 563.

Cecchi, A. S. (2014). Cognitive penetration, perceptual learning and neural plasticity. *Dialectica*, 68(1), 63–95.

Chase, W. G. and Simon, H. A. (1973a). The mind's eye in chess. In *Visual Information Processing*, edited by W. G. Chase, 215–81. New York, NY: Academic Press.

Chase, W. G. and Simon, H. A. (1973b). Perception in chess. *Cognitive Psychology*, 4, 55–81.

Chen, D., Yee Eng, H., and Jiang, Y. (2006). Visual working memory for trained and novel polygons. *Visual Cognition*, 14(1), 37–54.

Chudnoff, E. (2018). The epistemic significance of perceptual learning. *Inquiry*, 61(5–6), 520–42.

Chudnoff, E. (2020). Experience and epistemic structure: can cognitive penetration result in epistemic downgrade? In *Inference and Consciousness*, edited by A. Nes with T. Chan, 255–75. New York, NY: Routledge.

Cibelli, E., Xu, Y., Austerweil, J. L., Griffiths. T. L., and Regier, T. (2016). The Sapir–Whorf hypothesis and probabilistic inference: evidence from the domain of color. *PLoS One*, 11(7), e0158725.

Cohen, J. (2006). Color and perceptual variation revisited: unknown facts, alien modalities, and perfect psychosemantics. *Dialectica*, 60(3), 307–19.

Cohen, J. (2009). *The Red and the Real*. Oxford: Oxford University Press.

Curby, K. M. and Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin and Review*, 14(4), 620–8.

Curby, K. M., Glazek, K., and Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 94–107.

Dall, J. O., Wang, Y., Cai, X., Chan, R. C., and Sørensen, T. A. (2021). Visual short-term memory and attention: an investigation of familiarity and stroke count in Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(2), 282–94.

Dall, J. O., Watanabe, K., and Sørensen, T. A. (2016). Category specific knowledge modulates capacity limitations of visual short-term memory. In *2016 8th International Conference on Knowledge and Smart Technology (KST)*. IEEE, 275–80.

de Groot, A. D. and Gobet, F. (1996). *Perception and Memory in Chess*. Assen: Van Gorcum.

Elias, L. J., Saucier, D. M., Hardie, C., and Sarty, G. E. (2003). Dissociating semantic and perceptual components of synaesthesia: behavioural and functional neuroanatomical investigations. *Cognitive Brain Research*, 16(2), 232–7.

Elmera, S., Greber, M., Pushparaj, A., Kühnis, J., and Jäncke, L. (2017). Faster native vowel discrimination learning in musicians is mediated by an optimization of mnemonic functions. *Neuropsychologia*, 104, 64–75.

Firestone, C. and Scholl, B. J. (2016). Cognition does not affect perception: evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences*, 39, e229.

Gibson, J. J. and Gibson, E. J. (1955). Perceptual learning: differentiation or enrichment? *Psychological Review*, 62(1), 32–41.

Gilbert, C. D. and Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, 54, 677–96.

Gobet, F. and Simon, H. A. (1996). Templates in chess memory: a mechanism for recalling several boards. *Cognitive Psychology*, 31, 1–40.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585–612.

Green, E. J. (2019). A theory of perceptual objects. *Philosophy and Phenomenological Research*, 99(3), 663–93.

Grossenbacher, P. G. and Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Sciences*, 5(1), 36–41.

Hebb, D. O. (1949). *The Organization of Behavior: A Psychological Theory*. New York: Wiley.

Holt-Hansen, K. (1968). Taste and pitch. *Perceptual and Motor Skills*, 27(1), 59–68.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–54.

Jackson, M. C. and Raymond, J. E. (2005). Visual working memory: capacity is dependent on perceived, not physical, stimulus complexity. *Journal of Vision*, 5, 621.

Keane, D. (2017). Philosopher David Chalmers on consciousness, the hard problem and the nature of reality. ABC News, 6 July 2017. Available from: https://www.abc.net.au/news/2017-07-07/david-chalmers-and-the-puzzle-of-consciousness/8679884

Luck, S. J. and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–81.

Mannix, T. K. and Sørensen, T. A. (2019). Do we experience the same world? *Psykologi Information*, 47(1), 22–7.

Mannix, T. K. and Sørensen, T. A. (2021). Colours and category learning: implications for grapheme–colour synaesthesia. In *Sálubót: Afmælisrit til heiðurs Jörgen L. Pind*, edited by Á. Kristiánsson, H. M. Sigurdardottir, and K. Arnason, 221–30. Reykjavik: University of Iceland Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.

Mondloch, C. J. and Maurer, D. (2004). Do small white balls squeak? Pitchobject correspondences in young children. *Cognitive, Affective, and Behavioral Neuroscience*, 4, 133–6.

Nunn, J. A., Gregory, L. J., Brammer, M., *et al.* (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience*, 5, 371–5.

Pylyshyn, Z. W. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341–423.

Quilty-Dunn, J. (2020). Perceptual pluralism. *Noûs*, 54(4), 807–38.

Ransom, M. (2020). Attentional weighting in perceptual learning. *Journal of Consciousness Studies*, 27(7–8), 236–48.

Regier, T. and Xu, Y. (2017). The Sapir–Whorf hypothesis and inference under uncertainty. *WIREs Cognitive Science*, 8, e1440.

Reingold, E. M., Charness, N., Schultetus, R. S., and Stampe, D. M. (2001). Perceptual automaticity in expert chess players: parallel encoding of chess relations. *Psychonomic Bulletin and Review*, 8(3), 504–10.

Rosch, E. (1978). Principles of categorization. In *Cognition and Categorization*, edited by E. Rosch and B. B. Lloyd, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. New York, NY: Harcourt Brace.

Siegel, S. (2017). *The Rationality of Perception*. Oxford: Oxford University Press.

Simner, J. and Bain, A. E. (2013). A longitudinal study of grapheme–color synesthesia in childhood: 6/7 years to 10/11 years. *Frontiers in Human Neuroscience*, 7, 603.

Simner, J., Ward, J., Lanz, M., *et al.* (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology*, 22(8):1069–85.

Sørensen, T. A. (2019). Synaesthesia: atypical sensory experiences. In *Sense Me*, 142–51. Kolding: Trapholt.

Sørensen, T. A. and Kyllingsbæk, S. (2012). Short-term storage capacity for visual objects depends on expertise. *Acta Psychologica*, 140(2), 158–63.

Sørensen, T. A., Nordfang, M., and Ásgeirsson, Á. (2016). Gradual consolidation of synesthesia during adolescence: a case study. *Journal of Vision*, 16(12), 460.

Sørensen, T. A. and Overgaard, M. (2018). Prosopagnosia or prosopdysgnosia: facing up to a change of concepts. 10th International Conference on Knowledge and Smart Technology: Cybernetics in the Next Decades, KST 2018, IEEE Press (August 2018), 260–3.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, 74(11), 1–29.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–62.

Sutherland, N. (2017). *Cross-Modal Perceptual Learning: A Novel Shape Tasting Method for Sensory Discrimination of Wine*. [Thesis] Reed College.

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., and Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11), 4567–70.

Tilot, A. K., Kucera, K. S., Vino, A., Asher, J. E., Baron-Cohen, S., and Fisher, S. E. (2018). Rare variants in axonogenesis genes connect three families with sound–color synesthesia. *Proceedings of the National Academy of Sciences of the United States of America*, 115(12), 3168–73.

Tomson, S. N., Avidan, N., Lee, K., *et al*. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behavioral and Brain Research*, 223, 48–52.

Ward, J. (2013). Synesthesia. *Annual Review of Psychology*, 64(1), 49–75.

Wenger, M. J. and Rhoten, S. E. (2020). Perceptual learning produces perceptual objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(3), 455–75.

Whorf, B. L. (1956). Language, thought, and reality. In edited by J. B. Carroll. Cambridge, MA: MIT Press.

Witthoft, N. and Winawer, J. (2013). Learning, memory, and synesthesia. *Psychological Science*, 24(3), 258–65.

Witthoft, N., Winawer, J., and Eagleman, D. M. (2015). Prevalence of learned grapheme–color pairings in a large online sample of synesthetes. *PLoS One*, 10(3), e0118996.

Xie, W. and Zhang, W. (2017). Familiarity increases the number of remembered Pokémon in visual short-term memory. *Memory and Cognition*, 45(4), 677–89.