



# A review on various explanations of Ponzo-like illusions

Gizem Y. Yildiz<sup>1</sup> · Irene Sperandio<sup>2</sup> · Christine Kettle<sup>3</sup> · Philippe A. Chouinard<sup>1</sup>

Accepted: 31 August 2021 / Published online: 6 October 2021  
© The Psychonomic Society, Inc. 2021

## Abstract

This article reviews theoretical and empirical arguments for and against various theories that explain the classic Ponzo illusion and its variants from two different viewpoints concerning the role of perceived depth in size distortions. The first viewpoint argues that all Ponzo-like illusions are driven by perceived depth. The second viewpoint argues that the classic Ponzo illusion is unrelated to depth perception. This review will give special focus to the first viewpoint and consists of three sections. In the first section, the role of the number of pictorial depth cues and previous experience in the strength of all Ponzo-like illusions are discussed. In the second section, we contrast the first viewpoint against the theories that explain the classic Ponzo illusion with mechanisms that are unrelated to depth perception. In the last section, we propose a Bayesian-motivated reconceptualization of Richard Gregory's misapplied size constancy theory that explains Ponzo-variant illusions in terms of prior information and prediction errors. The new account explains why some studies have provided inconsistent evidence for misapplied size constancy theory.

**Keywords** Visual perception · Perceptual illusions · Ponzo illusion

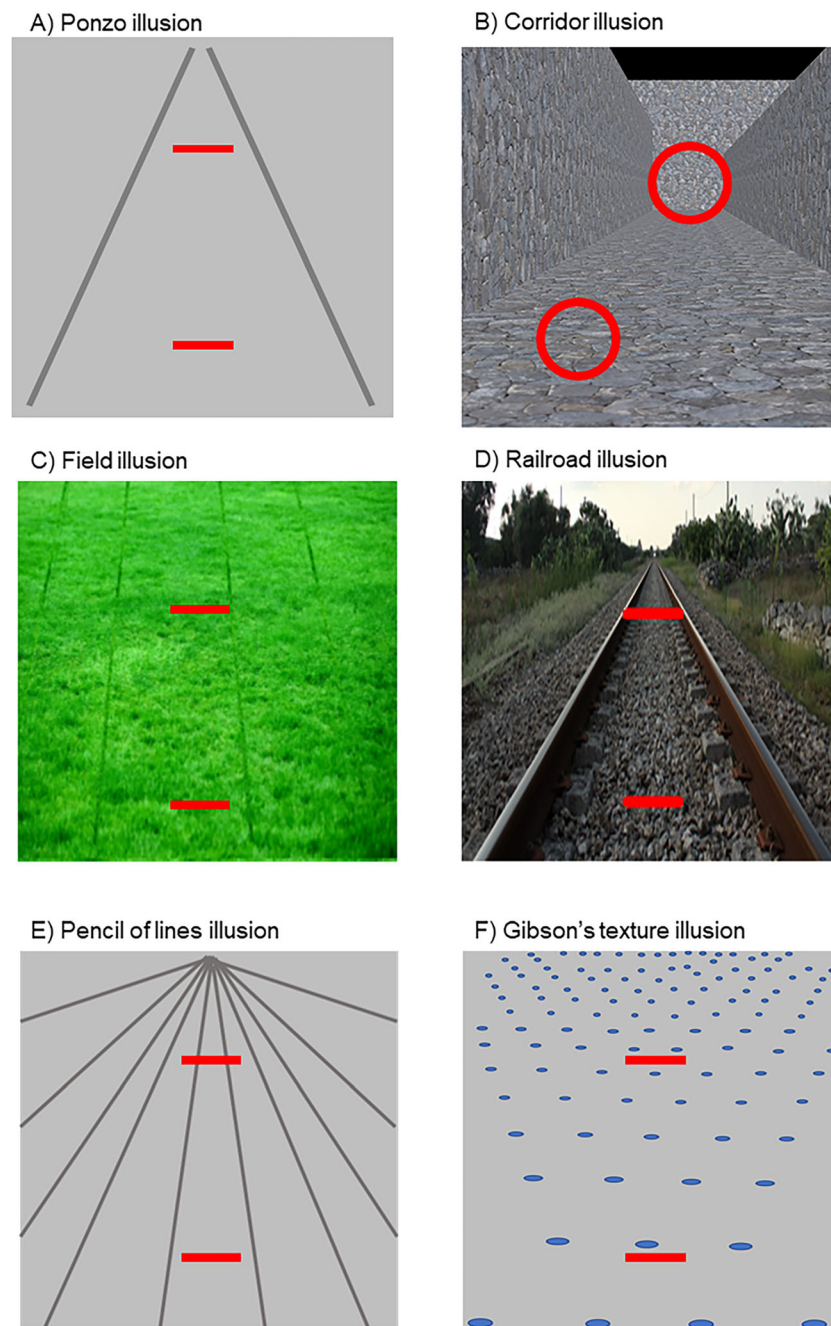
To interact with objects in the environment, we need to estimate their physical properties (e.g., size, shape, location, orientation) from the images on the retina. The objects' retinal images are continually changing when we see them under different viewing conditions. For example, objects that move farther away subtend gradually smaller visual angles on the retina. Yet we do not see them shrinking in size but rather as moving farther away from us. Remarkable distortions are evident when we see flat pictures of stimuli as two-dimensional(2D) representations of three-dimensional(3D)real-world scenes (Howe et al., 2006; Howe & Purves, 2005). For example, in the classic Ponzo illusion, two converging lines emulate a vanishing point in the upper section of the background image, similar to 2D representations

of sidewalls of a corridor in real-world(see Fig. 1a)(Ponzo, 1910).<sup>1</sup> When two physically identical red lines are placed at locations where converging lines of the classic Ponzo illusion signal different depth, they appear to be different from each other. Specifically, the bottom stimulus, placed at a location where contextual converging lines signal closer distance, appears to be smaller than the top stimulus, placed at a location where contextual converging lines signal greater depth. Similar effects are observable in one of the Ponzo-variant illusions shown in Fig. 1b. This illusion is called the corridor illusion. The classic Ponzo compared with the corridor background usually produces a weaker but still significant perceptual effect. Although these two illusions have attracted researchers' attention for more than a hundred years, there is no agreement on how to explain them. In reviewing the literature, it became apparent to us that the underlying mechanisms of the Ponzo-variant illusions and the classic Ponzo illusion

✉ Philippe A. Chouinard  
p.chouinard@latrobe.edu.au

<sup>1</sup> Department of Psychology and Counselling, School of Psychology and Public Health, La Trobe University, Melbourne, Victoria 3086, Australia  
<sup>2</sup> Department of Psychology and Cognitive Science, University of Trento, Rovereto, TN, Italy  
<sup>3</sup> Department of Pharmacy and Biomedical Sciences, School of Molecular Sciences, La Trobe University, Melbourne, Victoria 3086, Australia

<sup>1</sup> In the classic Ponzo illusion, two physically identical stimuli appear different from each other when they are placed over a background with converging contextual lines (Fig. 1a). A similar effect is observable in variant illusions in which a number of pictorial depth cues signal different depths at distinct sections of the background (e.g., corridor, field, railroad, pencil-of-lines illusions) (Fig. 1b–e). All these illusions will be referred to as either Ponzo-like or Ponzo-variant illusions throughout the paper.



**Fig. 1** Ponzo-like illusions. **a** The Ponzo illusion. The two physically identical lines are placed over the Ponzo background. Nonetheless, the top stimulus is seen as being bigger than the bottom one. The corridor (**b**), field (**c**), and railroad (**d**) illusions. When two physically identical stimuli are placed over the background, the top stimuli are seen as being bigger than the bottom stimuli. Similar background images were used in Leibowitz et al.'s (1969) study. **e** The pencil of lines illusion. The two physically identical lines are placed over the background with multiple linear perspective cues. Nevertheless, the top line is seen as longer than

the bottom one. A similar background was used in Leibowitz and Judisch's (1967) study. **f** Gibson's texture gradient pattern illusion. The two physically identical lines superimposed over the background with texture gradients have the same retinal image size. Fineman and Carlson (1973) tested the field illusion using Gibson's texture gradient pattern. The background produced no illusion when there is no linear perspective on how the texture patterns were arranged. (Color figure online)

are not entirely the same. These differences will be highlighted in this review.

This article aims to summarize and compare various explanations of the classic Ponzo and Ponzo-variant illusions,

giving particular emphasis to the misapplied size constancy theory, one of the most influential theories of visual illusions in the literature, which was proposed by Richard Gregory (1963, 1968, 1998) more than 50 years ago. The present

review is organized as follows. In the first section, empirical arguments for and against the misapplied size constancy theory will be reviewed. The role of the number of pictorial depth cues and previous experience in the strength of all Ponzo-like illusions will be discussed. In the second section, theoretical and empirical arguments for and against the theories that explain the classic Ponzo illusion with mechanisms that are unrelated to depth perception will be reviewed. Specifically, the contour-proximity (Fisher, 1968a, 1969, 1970, 1973), pool-and-store (Girgus & Coren, 1982), assimilation (Pressey, 1974b; Pressey et al., 1971; Pressey & Epp, 1992), and tilt constancy (Prinzmetal & Beck, 2001; Prinzmetal et al., 2001) theories will be considered and contrasted against the misapplied size constancy theory. In the third section, we will propose a reconceptualization of the misapplied size constancy theory within a Bayesian-framework. Hence, we will explain the Ponzo-like illusions in terms of prior information and prediction errors. This Bayesian interpretation will help us explain why some of the studies reviewed in the first two sections have provided evidence that goes against Gregory's account in an attempt to reconcile discrepancies in the literature.

## Misapplied size constancy theory

In the misapplied size constancy theory, Richard Gregory (1963, 1968, 1998) explained the size distortions that are experienced in the Ponzo-like illusions by means of inappropriate perceptual rescaling mechanisms. Perceptual rescaling mechanisms are normally helpful for estimating the size of 3D objects placed at different distances in the real 3D world. For example, two physically identical objects placed at different distances subtend different visual angles on the retina. Yet, we do not see them as having different sizes. This phenomenon is called size constancy (for a review, see Sperandio & Chouinard, 2015). Perceptual rescaling mechanisms can explain how we achieve size constancy. Namely, the brain uses binocular (e.g., binocular disparity, vergence angle) and monocular (the so-called pictorial depth cues, e.g., linear perspective cues and textures) sources of depth cues to estimate the distance between the eyes and the objects in the environment. After depth information is extracted, the objects' sizes are rescaled according to their perceived distance. This perceptual rescaling allows us to see the world as coherent and stable.

According to the misapplied size constancy theory, perceptual rescaling is inappropriate when it occurs in instances where the *apparent* rather than the *real* depth is processed by the brain. In other words, when the pictorial depth cues depicted in the visual illusions trigger perceptual rescaling mechanisms in flat pictures, inappropriate perceptual rescaling produces size distortions. Because many other and arguably more reliable sources of depth cues (e.g., binocular

disparity, lens accommodation, and vergence angle) are available to the visual system to support the perceptual rescaling of 3D real-world scenes (Dobias et al., 2016), size rescaling mechanisms have a greater effect on 3D real-world scenes than on 2D flat images with pictorial depth cues (Leibowitz et al., 1969; Sperandio et al., 2012). Given that binocular sources of depth cues, such as binocular disparity and vergence angle indicate that the stimuli placed at locations where pictorial depth cues signal different depths are, in fact, at equidistant locations from the eyes, the strength of the illusion has been argued to decrease under binocular compared to monocular viewing conditions (Yonas et al., 2002, but also see, Cretienoud et al., 2021).

Earlier studies have provided support for misapplied size constancy theory by showing that the magnitude of the illusion changes as a function of the number of available pictorial depth cues (Brislin, 1974; Fineman, 1981; Leibowitz et al., 1969; Yildiz et al., 2019, 2021a, 2021b) and frequent exposure to perspective cues such as corners and parallel lines in environments with city blocks, rectangular buildings, and street patterns (Brislin, 1974; Leibowitz et al., 1969; Leibowitz & Judisch, 1967; Leibowitz & Pick, 1972; Wagner, 1977). For example, Leibowitz et al. (1969) demonstrated that both the classic Ponzo (see Fig. 1a) and field (Fig. 1c) backgrounds produced a significant but weaker illusion than the railroad background (Fig. 1d), which conceivably included both the Ponzo and field backgrounds. In the same study, Leibowitz et al. (1969) compared the magnitude of the illusions in students from Guam and Pennsylvania in the United States. The authors reasoned that if the strength of the illusions was modulated as a function of the exposure to pictorial depth cues, then these illusions would be stronger in Pennsylvanian students, who have had greater exposure to the pictorial depth cues in their daily life. Supporting this line of reasoning, frequent exposure to linear perspective cues and longer vistas increased the strength of the field and railroad illusions. These findings were in line with Leibowitz and Judisch's (1967) study demonstrating how the developmental trajectory of a Ponzo-variant illusion (specifically, the pencil-of-lines illusion) was similar to the developmental trajectory of size constancy in typically developing children (Zeigler & Leibowitz, 1957). The role of the number of pictorial depth cues and previous experience in the strength of all Ponzo-like illusions is reviewed in the next subsections.

However, the misapplied size constancy explanation of the Ponzo illusion has been challenged in three major ways. First, the background images that were used in Leibowitz et al.'s (1969) study have been criticized as being poor demonstrations of systematic pictorial depth cue removal (see Figs. 1a, c–d; Fineman & Carlson, 1973). Specifically, their field background has been criticized for including the linear perspective

cues of the classic Ponzo illusion (Fineman & Carlson, 1973). Contrary to the Leibowitz et al.'s (1969) findings, Fineman and Carlson (1973) demonstrated that the field background does not produce an illusion when there is no linear perspective in how the texture patterns are arranged (see Fig. 1f). Because the classic Ponzo but not the field background produced an illusion, Fineman and Carlson concluded that the conceptual understanding of depth might not drive the classic Ponzo illusion. Second, cross-cultural studies revealed evidence that frequent exposure to pictorial depth cues affects the magnitudes of the field (Fig. 1c) and railroad (Fig. 1d) illusions (Brislin, 1974; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Wagner, 1977), while they have little or no effect on the magnitude of the classic Ponzo illusion (see Fig. 1a; Brislin, 1974; Jahoda & McGurk, 1974; Leibowitz et al., 1969; McGurk & Jahoda, 1975; Segall et al., 1966). Third, Brislin (1974) demonstrated that the strength of the context-rich field (Fig. 1c) and railroad (Fig. 1d) illusions, but not the classic Ponzo illusion (Fig. 1a), increased with age in typically developing children. Because the strength of the classic Ponzo illusion does not vary as a function of previous experience, these findings suggest that the conceptual understanding of depth might not drive the classic Ponzo illusion. The influences of the number of pictorial depth cues and previous experience on the strength of all Ponzo-like illusions are discussed in the following subsections.

### The magnitude of the illusion as a function of the availability of pictorial depth cues

Many have argued that if all Ponzo-like illusions are driven by perceived depth, then the strength of the illusions would increase with an increase in the amount of pictorial depth cues (Brislin, 1974; Fineman, 1981; Leibowitz et al., 1969). Indeed, it is conceivable that the converging lines of the classic Ponzo background in Fig. 2a are a poor demonstration of how linear lines recede into depth as compared with the converging lines of a corridor background displayed in the same figure. Supporting this line of reasoning, Chevrier and Delorme (1983) reported that perceived depth increased when the stimuli were presented over a background with linear perspective cues and textures compared with a background with only linear perspective cues (Fig. 2d).

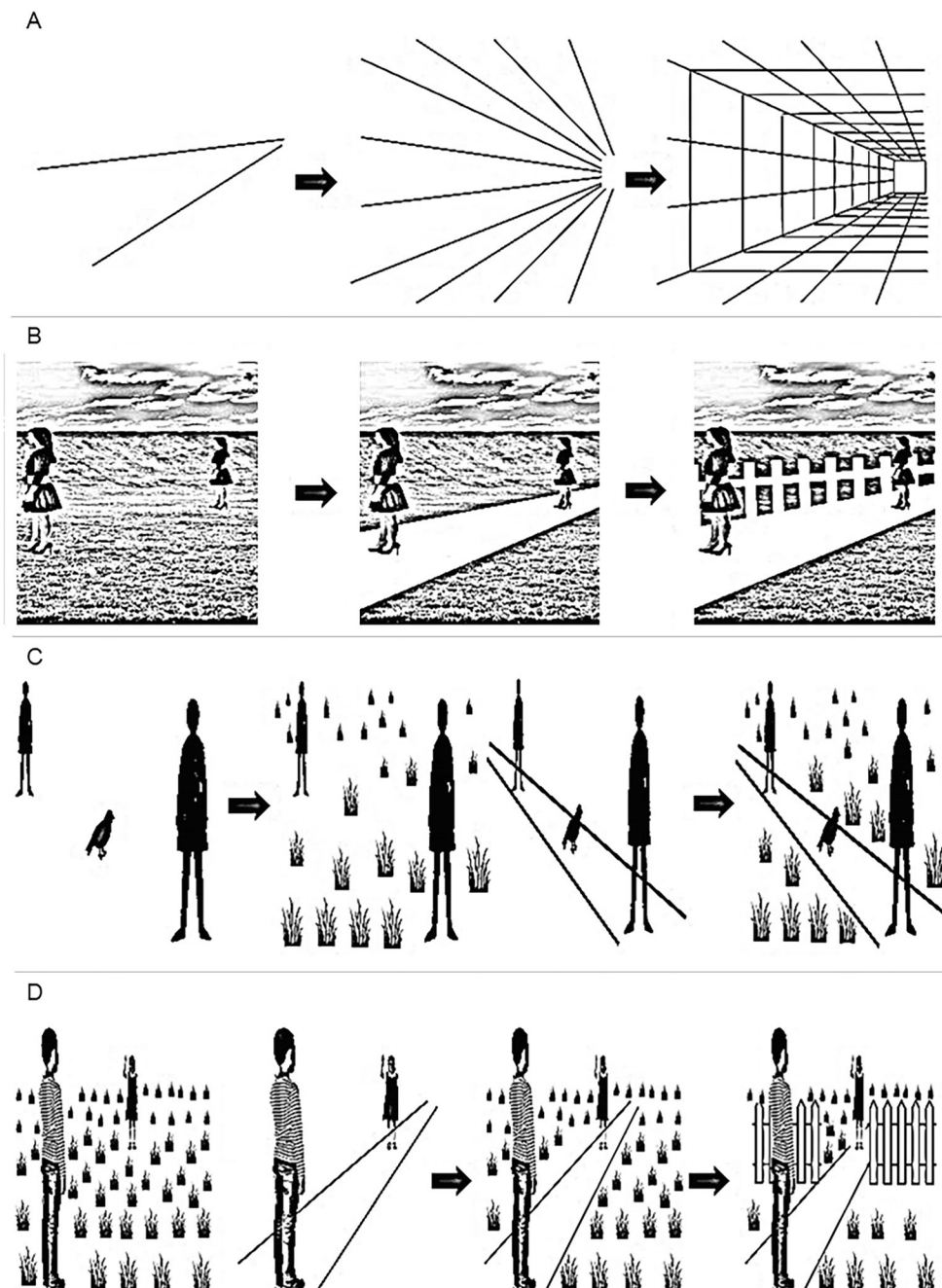
Previous studies have demonstrated that the magnitude of the illusion changes as a function of the number of pictorial depth cues that are added to the illusion display (Brislin, 1974; Cretenoud et al., 2020; Fineman, 1981; Jahoda & McGurk, 1974; Kilbride & Leibowitz, 1975; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Wagner, 1977). In some studies, backgrounds with pictorial depth cues were composed of simple line drawings (i.e., see Fig. 2; Chevrier & Delorme, 1983; Fineman, 1981; Jahoda & McGurk, 1974; McGurk & Jahoda,

1975; Wohlwill, 1962). Therefore, it is possible to question the ecological validity of the background images used in these instances. Perhaps because there was no better option, some researchers continued to use Leibowitz et al.'s (1969) background images (Fig. 1a and Fig. 1c–d), which were criticized as being poor demonstrations of systematic pictorial depth cue removal (Fineman & Carlson, 1973), to test the role of previous experience in the effects of the number of pictorial depth cues in perceptual rescaling (Brislin, 1974; Kilbride & Leibowitz, 1975; Leibowitz & Pick, 1972; Wagner, 1977).

Recently, we did a better job at systematically removing the texture gradients and linear perspective cues from a corridor background image (Fig. 3a–c) and tested the role of these two pictorial depth cues in perceptual size rescaling (Yildiz et al., 2019). In line with studies showing how the strength of the Ponzo illusion increased with an increase in the number of pictorial depth cues, the strength of the corridor illusion was greater when the rings were presented on the background with texture gradients and linear perspective cues. This was particularly evident when the strength of the illusion produced by texture gradients (Fig. 3b) was compared with the illusion produced by the background with both texture gradients and linear perspective cues (Fig. 3c). Surprisingly, there were no significant differences between the magnitudes of the corridor illusions produced by the backgrounds with linear perspective cues (see Fig. 3a and c; Yildiz et al., 2019). Taken together, the former but not the latter finding provided evidence for misapplied size constancy theory.

In another study, we tested whether the perceptual rescaling of stimuli with linear perspective cues and textures occurs at earlier or later stages of visual analysis using an interocular transfer paradigm (Yildiz et al., 2021b; Fig. 3d). We reasoned that if perceptual size rescaling takes place at later stages of visual analysis, then the influence of background would transfer from one eye to the other, so that there would not be a significant difference between the strength of the illusions produced under the monocular (Fig. 3e) and dichoptic (Fig. 3f) viewing conditions. Our results revealed that the background with linear perspective cues produced a greater illusion under the monocular viewing condition, while the background with texture gradients produced an equal amount of illusion under the monocular and dichoptic viewing conditions. When we repeated the experiment using a classic Ponzo background, our findings revealed that contextual converging lines produced an equal amount of illusion under the monocular and dichoptic viewing conditions. Taken together, these findings suggested to us that the visual system could extract depth information effortlessly with monocular neural populations from the background with linear perspective cues in the corridor illusion. Contrarily, the extraction of depth information from the background with texture gradients and the classic Ponzo background requires the involvement of the binocular neural populations in higher-level cortical areas. As Gregory's misapplied size constancy theory





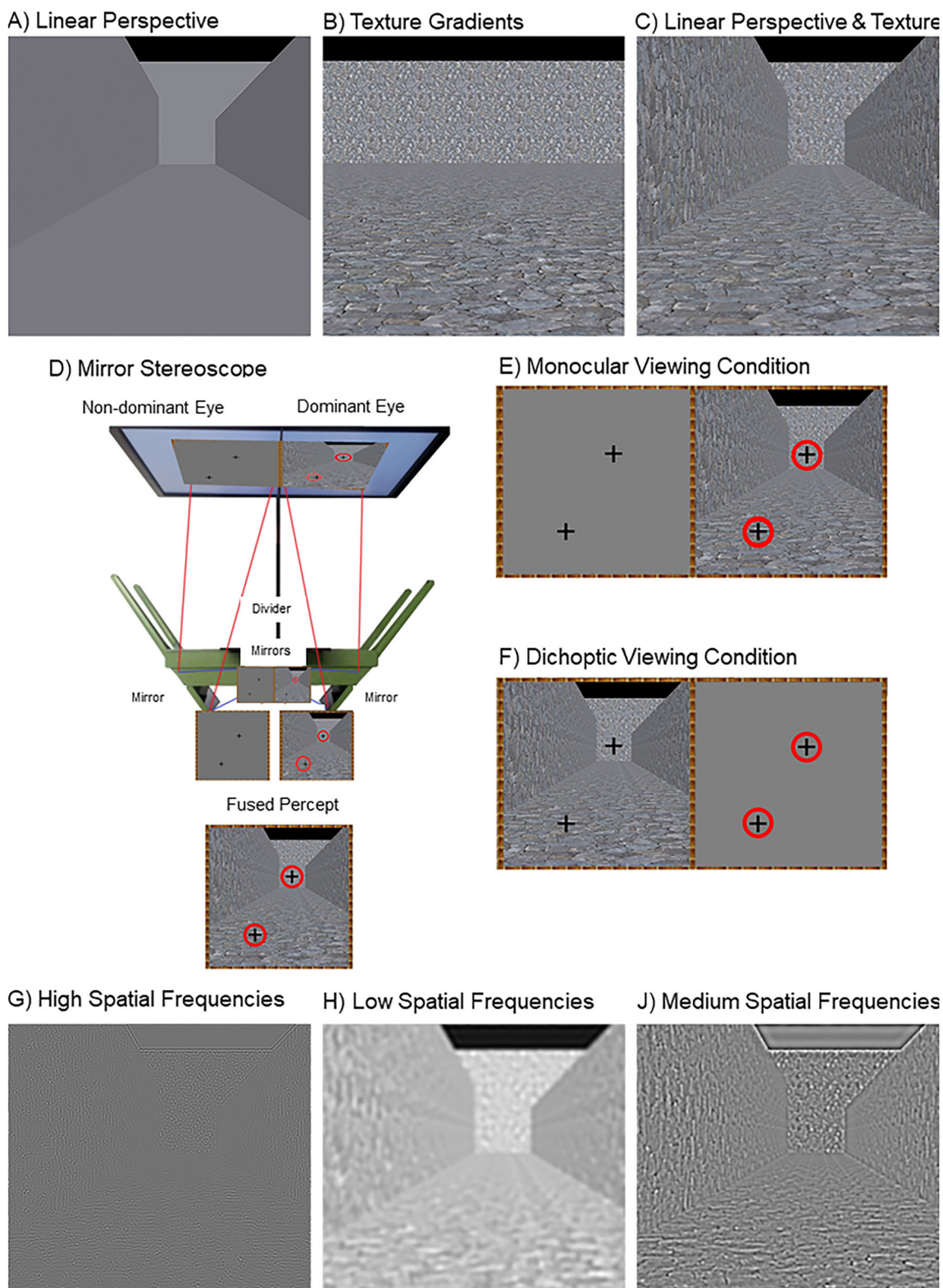
**Fig. 2** Background images used to test how perceptual size rescaling of stimuli changes with the availability of pictorial depth cues. **a** Fineman (1981) tested the effect of linear perspective cues on perceptual size rescaling mechanisms by adding them in a deformed version of the classic Ponzo background. **b** Jahoda and McGurk (1974) tested the effect of previous experience on the Ponzo-like illusions by adding pictorial depth cues (a road and a fence) in a control background. **c** McGurk and Jahoda

(1975) tested the effect of previous experience on the Ponzo-like illusions by adding pictorial depth cues (texture gradients and linear perspective cues) in a control background. **d** Chevrier and Delorme (1983) tested the effect of previous experience on the Ponzo-like illusions by adding pictorial depth cues (texture gradients and linear perspective cues) in a background

predicts no difference for the underlying mechanisms of the textures and linear perspective cues in the classic Ponzo and corridor illusions, these findings provided inconsistent evidence for his original theory.

In another study, we systematically removed the high, medium, and low spatial frequencies from the corridor

background with texture gradients and linear perspective cues (see Fig. 3c; Yildiz et al., 2021a). It was easier to see the edges of the texture gradients in the background with high spatial frequencies (Fig. 3g), while it was easier to see the linear perspective cues in the background with low spatial frequencies (Fig. 3h). The backgrounds with either the high (Fig. 3g),



low (Fig. 3h), and medium (Fig. 3i) spatial frequencies produced a significant but weaker illusion than the original background with all spatial frequencies (Fig. 3c). Notably, there were no differences among the magnitudes of the corridor

illusions produced by the backgrounds with the high (Fig. 3g), low (Fig. 3h), and medium (Fig. 3i) spatial frequencies. Since both the texture gradients and linear perspective cues were visible in the background with medium spatial

◀ **Fig. 3** Experimental stimuli used in our previous studies. **a–c** Backgrounds used in Yildiz et al. (2019, 2021a). **a** A corridor background display of a hallway with walls (linear perspective cues). **b** A corridor background display of a hallway with stones (texture gradients). **c** A corridor background display of a hallway with walls (linear perspective cues) and stones (texture gradients). **d–f** Apparatus and experimental stimuli used in Yildiz et al.'s (2021b) study. **d** A mirror stereoscope was used to test the interocular transfer effects of texture gradients and linear perspective cues. **e** Under the monocular viewing condition, the background images with pictorial depth cues and target rings were presented together to the dominant eye. **f** Under the dichoptic viewing condition, the background images with pictorial depth cues and target rings were presented separately to the non-dominant and dominant eyes, respectively. **g–i** Backgrounds used in Yildiz et al. (2021a). **g** The corridor background image with high spatial frequencies. **h** The corridor background image with low spatial frequencies. **i** The corridor background image with medium spatial frequencies. (Color figure online)

frequencies, the latter finding provided evidence against misapplied size constancy theory.

Although many have argued that Gregory's misapplied size constancy theory predicts an increase in the illusion's magnitude with an increase in the number of pictorial depth cues, some of our findings contradict this prediction (Yildiz et al., 2019, 2021a, 2021b). Additionally, Gregory's original theory predicts no difference for the underlying mechanisms of the textures and linear perspective cues in the Ponzo and corridor illusions. Contrary to this hypothesis, we found differences in the underlying mechanisms of textures and linear perspective cues in the classic Ponzo and corridor illusions (Yildiz et al., 2021b). To reconcile these discrepancies, in the third section of our review paper, we propose a Bayesian-motivated reconceptualization of misapplied size constancy that explains all Ponzo-variant illusions in terms of prior information and prediction errors. For a reader who is familiar with Gregory's works, it is interesting to see how Gregory's views developed over the years. For example, between 2005 and 2006, Gregory (2005, 2006a, 2006b) published a series of editorial essays on the *Bayes window*. In these essays, Gregory argued that the brain extracts depth information from the 2D backgrounds covering pictorial depth cues with prior knowledge in consideration. Our Bayesian-motivated reconceptualization of his theory is based on these essays.

### The magnitude of the illusion as a function of previous experience with pictorial depth cues

Many have argued that if the Ponzo-like illusions are related to perceived depth, then the strength of the illusions should change depending on previous experience with pictorial depth cues (Brislin, 1974; Kilbride & Leibowitz, 1975; Leibowitz et al., 1969; Wagner, 1977). It is conceivable that a person

who has more experience in computing object size at different distances would implicitly apply this information to extract depth from 2D pictures more effectively than a person with less experience.

Indeed, previous studies have shown that perceptual responses vary depending on previous experience (Davidoff et al., 1999; de Fockert et al., 2007; Deregowski, 1989, 2017; Doherty et al., 2008; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Miller, 1973). Some have explained differences in visual perception by differences in how the visual system adapts to a variety of environments to interpret physical information using previous experience (Brislin, 1974; Kilbride & Leibowitz, 1975; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Wagner, 1977). The effect of previous experience on the Ponzo-like illusions has been studied by focusing on culture- and age-related changes in the strength of the illusions.

### The effects of culture on illusion magnitude

The effects of culture on the strength of Ponzo-like illusions have been studied extensively by focusing on the effects of exposure to different surrounding environments on illusion magnitude. Several cross-cultural studies have sought to investigate how the strength of Ponzo-like illusions changes with frequent exposure to real-life pictorial depth cues in the surrounding environment and their 2D representations in pictures. Table 1 provides a summary of studies examining the effects of culture on the magnitude of the Ponzo-like illusions and depth perception. In line with Gregory's misapplied size constancy theory, Segall et al. (1966) demonstrated that living in an environment with city-blocks, rectangular buildings, and street patterns increases the strength of the Müller-Lyer (see Fig. 4a; Müller-Lyer, 1889) and Sander's parallelogram (see Fig. 4b; Sander, 1926) illusions. Contrary to the misapplied size constancy theory's prediction, cultural differences had no impact on the Ponzo illusion (Fig. 4c). Because a deformed version of the Ponzo background was used to measure the perceived size differences between the top and bottom horizontal lines, Deregowski (2017) suggested that unusualness of the background image may have hindered the effects of cultural differences on the strength of the Ponzo illusion.

Contrary to Deregowski's interpretation, many have demonstrated that Segall et al.'s findings could not be explained by their choice of the Ponzo background (Brislin, 1974; Jahoda & McGurk, 1974; Leibowitz et al., 1969; McGurk & Jahoda, 1975). For example, Brislin (1974) reported that a group of participants living in environments with little or no example of pictorial depth cues might still experience as strong a classic Ponzo illusion (Fig. 1a) as participants living in environments with many examples of pictorial depth cues.

There is reason to assume that cross-cultural differences in the magnitude of the Ponzo-like illusions are related to differences in the perception of depth from 2D flat images with



**Table 1** Summary of previous studies examining the effect of culture on the magnitude of the Ponzo-like illusions and perceived depth

Author(s), year	Sample	N	Stimuli	Design	Results
Hudson (1960)	School-going and non-school-going participants from the Union of South Africa	562	2D flat images with linear perspective, relative size, and interposition pictorial depth cues (e.g., Fig. 4d)	For Fig. 4d, participants were asked to answer the following questions: (1) What do you see? (2) What is the man doing? (3) Which is nearer the man: elephant or antelope? Based on the last question, participants were classified as either 2D or 3D perceivers. The first two questions were used to detect the role of correct identification in the perceived depth.	Almost all school-going participants reported seeing that the animal, placed at a location where pictorial depth cues signalled little or no depth, was nearer the man. These participants were classified as 3D perceivers. Contrarily, almost all non-school-going participants reported seeing that the animal, placed at a location where pictorial depth cues signalled greater depth, was nearer the man. These participants were classified as 2D perceivers. Correct identification of pictorial depth cues had little influence on 3D perception.
Segall et al. (1966)	Three Western and 14 non-Western (i.e., African, Haunŋoo Philippines) samples	1,878	The Ponzo (Fig. 4c), Müller-Lyer (Fig. 4a), and Sander parallelogram (Fig. 4b) illusions	The participants' task was to indicate the longer line.	Western people experienced stronger Müller-Lyer and Sander's parallelogram illusions than the non-western people. Contrary to these findings that provided support for Gregory's misapplied size constancy theory, cultural differences had no impact on the strength of the Ponzo illusion.
Leibowitz and Pick (1972)	Ugandan origin students studying at Makerere University and Ugandan villagers	140	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) background images	Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line presented at the upper section of the background image.	The strength of the illusion increased depending on the number of available depth cues among Ugandan university students but not among villagers. Namely, for Ugandan villagers, none of the 2D backgrounds produced an illusory effect.
Leibowitz et al. (1969)	Students studying at the Pennsylvania State University and Guam origin students studying at the University of Guam	44	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) background images	Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line presented at the upper section of the background image.	Frequent exposure to linear perspective cues and longer vistas increased the strength of the field (Pennsylvanian students: ~21%, Guamanian students: ~12%) and railroad (Pennsylvanian students: 32%, Guamanian students: 18%) illusions among Pennsylvanian students who exposed more to the pictorial depth cues in real-world scenes. Contrarily, frequent exposure to linear perspective cues and longer vistas had little or no effect on the magnitude of the Ponzo illusion.
Brislin (1974)	Natives of Guam island and residents of Pennsylvania	200	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) background images	Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line presented at the upper	There were no differences between Pennsylvanian and Guamanian participants in the strength of the Ponzo illusion. Nonetheless, cultural differences affected how the strength of the illusion increased with an increase in the number of pictorial depth cues.



**Table 1** (continued)

Author(s), year	Sample	N	Stimuli	Design	Results
				section of the background image.	Namely, Pennsylvanian participants experienced stronger field and railroad illusions than Guamanian participants.
Wagner (1977)	Moroccan males	384	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) background images	<p>1) Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line presented at the upper section of the background image.</p> <p>2) Pictorial depth perception: Participants were asked to report what they have seen over the railroad backgrounds.</p>	Participants who reported to see a path, road, or track were classified as 3D perceivers, while participants who reported to see a ladder, hill, or tent were classified as 2D perceivers. 3D perceivers experienced stronger field and railroad illusions than the 2D perceivers. This classification had no effect on the perceived size of the lines presented over the Ponzo and control backgrounds. In the Ponzo illusion, neither urbanization nor schooling affected the perceived size differences. Conversely, both schooling, urbanization, and age affected the perceived size differences in the field and railroad illusions. Namely, the oldest schooled children living in an urban area experienced stronger railroad and field illusions.
Kilbride and Leibowitz (1975)	Ugandan villagers	82	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) background images	<p>1) Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line, presented at the upper section of the background image.</p> <p>2) Pictorial depth perception: A road image and the Ponzo background were presented to the participants. Participants were classified as 2D, 3D, or mixed perceivers based on their responses.</p>	Ugandan villagers were susceptible to the Ponzo illusion, but the strength of the illusion did not increase with the addition of pictorial depth cues among 2D perceivers. 3D perceivers experienced stronger field and railroad illusions than the 2D and mixed perceivers while there were no differences among 2D, mixed, and 3D perceivers in the Ponzo illusion. Notably, the strength of the illusion increased with the number of depth cues in only 3D perceivers. Additionally, the authors reported that minimizing flatness depth cues had little or no impact on the strength of the railroad, field, and Ponzo illusions.
Jahoda and McGurk (1974)	African, Chinese, and European children	167	The backgrounds with elevation, elevation + linear perspective (a road), elevation + rich perspective (a road with a fence) pictorial depth cues (Figure 2B)	3D model construction task: The participants were asked to construct a 3D model of the 2D picture using bigger or smaller wooden figures depicting an adult or a child, respectively.	Cultural differences had little or no impact on participants' performance. In terms of size accuracy, participants' performance increased with an increase in age and the number of pictorial depth cues in the 2D image. In terms of spatial arrangement accuracy, participants' performance increased with an increase in age,

**Table 1** (continued)

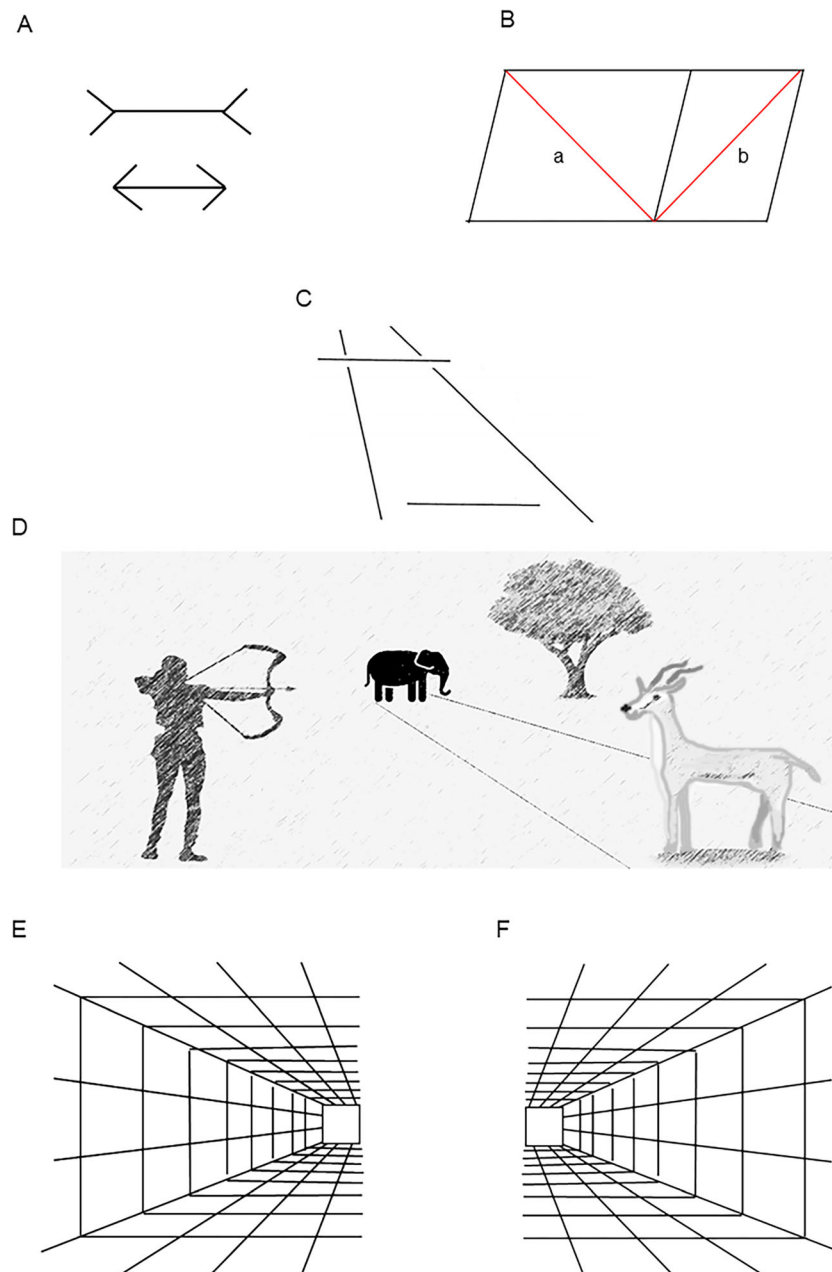
Author(s), year	Sample	<i>N</i>	Stimuli	Design	Results
McGurk and Jahoda (1975)	Children living in Scotland and Ghana	208	The backgrounds with elevation, elevation + linear perspective (a road), elevation + texture gradients (a field), and elevation + linear perspective + texture gradients (Fig. 2c)	3D model construction task: The participants were asked to construct a 3D model of the 2D picture using bigger or smaller wooden figures.	but it decreased in the presence of linear perspective depth cues. Scottish children performed better than Ghanaian children in both size and spatial arrangement tasks.
Rima et al. (2019)	French and Syrian participants	80	The rightward (Fig. 4e) and leftward (Fig. 4f) corridor images	Method of adjustment: The standard line was presented at the apex of the corridor. The participants were asked to match the size of the comparison line presented at the open end of the corridor to the standard line's size.	The illusion was stronger among French participants when the rightward (left foreground/right background) corridor image was presented. Contrarily, the illusion was stronger among Syrian participants when the leftward (right foreground/left background) corridor image was presented. The authors explained these findings with reading/writing habits.

pictorial depth cues. For example, a group of participants living in environments with many examples of pictorial depth cues might still experience a weaker or no Ponzo-like illusions if they have little or no previous experience with 2D representations of 3D scenes that contain pictorial depth cues (i.e., looking at pictures, photos, television; see Deregowski, 1989, for a review). Supporting and confirming this hypothesis, Leibowitz and Pick (1972) demonstrated that the strength of the illusion increased depending on the number of available depth cues in Ugandan Makerere University students with considerable experience seeing pictures, but not Ugandan villagers with less experience seeing pictures. Because it is possible to see many examples of pictorial depth cues with large terrains, rectangular farms, square homes, and tarmac roads in Uganda, this finding could not be explained by the frequency with which Ugandan villagers are exposed to pictorial depth cues. Thus, repeated exposure to pictures that depict 2D representations of conventional 3D pictorial depth cues can affect perceived stimulus size.

In line with this, both schooling (Deregowski, 1968; Hudson, 1960) and frequent passive exposure to 2D representations of 3D scenes (Mundy-Castle, 1966) were found to be important to perceiving depth from 2D flat surfaces with pictorial depth cues. For example, Hudson (1960) tested the role of schooling in perceiving depth from 2D flat images with pictorial depth cues. One of the 2D flat images with pictorial depth cues used in this study is displayed in Fig. 4d. Hudson (1960) reasoned that if the experience with 2D representations of 3D scenes affects the perception of pictorial depth, then the

participants who had never attended school would less likely perceive the elephant in Fig. 4d as being located further away than the hunter and antelope, despite the availability of obvious pictorial cues that indicate different depths, such as linear perspective and relative size. The participants who reported to see that the hunter was aiming at the elephant were classified as 2D perceivers while the participants who reported to see that the hunter was aiming at the antelope were classified as 3D perceivers. This classification revealed that the participants who had attended school were almost entirely 3D perceivers, while the participants who had never attended school were almost entirely 2D perceivers.

Hudson's (1960) classification method has been used as an alternative approach to test the role of cross-cultural differences in the magnitude of the Ponzo-like illusions. Many have argued that if the participants could not recognize pictorial depth cues in the 2D flat images, then they should not experience a stronger illusion with the addition of these depth cues (Kilbride & Leibowitz, 1975; Wagner, 1977). In agreement with this hypothesis, Kilbride and Leibowitz (1975) demonstrated that 3D perceivers experienced stronger field and railroad illusions than the 2D perceivers. Yet, there were no differences between 2D and 3D perceivers in the classic Ponzo illusion. Therefore, the strength of the illusion increased with the number of depth cues only in 3D perceivers. Additionally, Kilbride and Leibowitz (1975) reported that minimizing flatness depth cues had little or no impact on the strength of the railroad, field, and classic Ponzo illusions. This finding suggested to the authors that



**Fig. 4** Background images used to test the effect of culture on the magnitude of the illusions. **a** Müller-Lyer illusion. The two horizontal lines are physically identical. Yet the top line appears larger than the bottom line. A similar illusion was used in the Segall et al. (1966) study. **b** Sander's parallelogram illusion. The two red lines are physically identical in size. Yet we see line X longer than line Y. A similar illusion was

used in the Segall et al. (1966) study. **c** A deformed version of the classic Ponzon background. A similar background was used in the Segall et al. (1966) study. **d** A background with pictorial depth cues used in the Hudson (1960) study. The rightward **e** and leftward **f** corridor backgrounds used in the Rima et al. (2019) study

2D perceivers experienced weaker railroad and field illusions because they were unable to recognize pictorial depth cues in the scene.

Similarly, Wagner (1977) showed that 3D perceivers experienced stronger field (Fig. 1c) and railroad (Fig. 1d) illusions than 2D perceivers. This classification had little or no effect on the perceived size of the lines presented over the classic Ponzon (Fig. 1a) and control backgrounds. These findings suggested

to the author that magnitudes of the railroad and field illusions changed depending on the perceptual skills while the classic Ponzon illusion was insensitive to perceptual experience.

Additionally, Wagner (1977) tested the role of urbanization, age, and schooling in the strength of Ponzon-like illusions. In the classic Ponzon illusion, neither urbanization nor schooling affected perceived size differences. Conversely, schooling, urbanization, and age affected the perceived size



differences in the field and railroad illusions. Namely, the oldest schooled children living in an urban area experienced stronger field and railroad illusions. In a similar vein, Kilbride and Robbins (1968) showed that the amount of educational level predicted how well Indian participants extracted depth information from 2D representations of 3D scenes. Overall, these studies showed that cultural differences affect how the strength of the illusion increases with an increase in the number of pictorial depth cues.

Nonetheless, there are studies showing that cultural differences have no impact on the strength of Ponzo-like illusions (Jahoda & McGurk, 1974; McGurk & Jahoda, 1975). For example, Jahoda and McGurk (1974) asked African children living in rural villages, Chinese children living in either urban areas of Hong Kong or on Chinese riverboats, and European children living in Scotland to construct a 3D model of a 2D picture with pictorial depth cues using bigger or smaller wooden figures depicting an adult or a child, respectively. The authors reported that the strength of the illusion increased with an increase in age and the number of pictorial depth cues in the 2D image. Contrary to the findings of Leibowitz and his colleagues (1969), in their study, the strength of the illusion increased with an increase in the number of pictorial depth cues regardless of cultural differences. Based on these findings, the authors suggested that cultural differences may play a weaker role in perceptual rescaling of stimuli with the pictorial depth cues than what was thought before. Differences between the tasks, background images, and measurement methods might explain why there is a discrepancy between the findings by McGurk and Jahoda and the findings by Leibowitz and his colleagues.

The influence of reading and writing habits on perceptual responses between cultures has also been examined (Friedrich & Elias, 2014; Rima et al., 2019). For example, Rima et al. (2019) compared the magnitude of the rightward and leftward corridor illusion between French and Syrian participants, who read in different directions (Fig. 4e–f). Their results revealed that the illusion was stronger among French participants, whom use the rightward reading/writing system, when the rightward (left foreground/right background) corridor image (Fig. 4e) was presented. Contrarily, the illusion was stronger among Syrian participants, whom use the leftward reading/writing system, when the leftward (right foreground/left background) corridor image (Fig. 4f) was presented. The authors explained these findings in terms of reading and writing habits. They argued that French participants who were left-to-right readers experienced a stronger illusion with the rightward corridor because they organize elements of an image more easily while scanning pictures from left to right than the other direction. The same logic was used to explain why the Syrian participants had a stronger illusion for the other background configuration. These findings might be explained by how reading and writing habits influence how we direct the

focus and distribute our attention toward objects around our environment. In line with this, Chokron and De Agostini (2000) argued that people tend to direct their attention toward the side on which they begin reading.

Together, these results suggest that if participants have previous experience with how specific pictorial depth cues affect the size of objects in real life, then the magnitude of the illusion changes depending on the number of available pictorial depth cues. This is particularly evident when Leibowitz et al.'s backgrounds (Fig. 1a and Fig. 1c–d) were used (Brislin, 1974; Kilbride & Leibowitz, 1975; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Wagner, 1977). Those studies in which all background images were composed of simple line drawings (Fig. 2b–c; Jahoda & McGurk, 1974; McGurk & Jahoda, 1975) revealed that cultural differences have no impact on how the magnitude of the illusion changes depending on the number of available pictorial depth cues. Notably, in Leibowitz et al.'s studies, both the field (Fig. 1c) and railroad (Fig. 1d) backgrounds were full-tone pictures of real 3D scenes, while the classic Ponzo background (Fig. 1a) was composed of simple line drawings. This discrepancy might result in differences across cultural groups.

Additionally, the cross-cultural studies that we have reviewed here have also revealed that all Ponzo-like illusions are present across all cultures—even though some may experience a weaker illusion. Because having a default system that extracts depth information from linear perspective cues is vital to survival, evolutionarily programmed mechanisms might explain the presence of the illusions across cultures. Gregory's original misapplied size constancy theory predicts that the strength of all Ponzo-like illusions would increase with previous experience. Therefore, the finding showing that previous experience has no effect on the strength of the classic Ponzo illusion does not support such a theory. In the third section of this review paper, we propose a Bayesian-motivated reconceptualization of misapplied size constancy theory that explains the Ponzo-like illusions by means of prior information and prediction errors. This reconceptualization clarifies why some studies have provided inconsistent evidence for Gregory's misapplied size constancy theory.

### The effects of age on the magnitude of the illusion

It is conceivable that an adult with many years of experience in computing object size at different distances would implicitly apply this information to extract depth from 2D pictures more effectively than a young child whose 3D spatial skills are still developing and who has computed fewer object sizes using pictorial depth cues. According to the misapplied size constancy theory, age-dependent changes in the Ponzo-like illusions are likely related to (1) the sensitivity of the visual system to different pictorial depth cues, (2) the ability to perceive objects as having the same size despite changes in

viewing distance (i.e., size constancy), and (3) the ability to build 3D representations of the visual scenes from 2D pictures. Several developmental studies have sought to investigate how size constancy, sensitivity to pictorial depth cues, and the strength of the Ponzo-like illusions change with age. Table 2 provides a summary of the studies examining the effect of age on the magnitude of the Ponzo-like illusions and depth perception.

Previous studies demonstrated that infants are sensitive to the presence of some pictorial depth cues (Hemker & Kavsek, 2010; Yonas et al., 1978; Yonas et al., 2002). For example, Hemker and Kavsek (2010) reported that 7-month-old infants preferred to reach for the objects placed at a location where the linear perspective cues, but not the texture gradients, signalled greater depth. It could be speculated that our brains have evolutionarily been programmed to process linear perspective cues more effectively than texture gradients. Alternatively, because infants have reduced visual acuity (Maurer et al., 1999; Teller, 1997), it should not be surprising that they are less sensitive to the presence of texture gradients, which conceivably requires acuity for fine details, than linear perspective cues, which are still more easily detectable in the coarse images.

Although sensitivity to linear perspective cues might have an important innate component, how development leads to increased sensitivity to pictorial depth cues in infants remains unknown. Evidence from longitudinal studies suggests that perceptual experience plays an important role in the development of sensitivity to pictorial depth cues, including linear perspective (Yonas et al., 2002). Therefore, sensitivity to pictorial depth cues increases with age.

Studies investigating size constancy in infants have demonstrated that infants, including newborns, respond to physical, rather than the retinal size of objects at different distances (Granrud, 2006; Slater et al., 1990). For example, Granrud (2006) investigated size constancy for near objects in 4-month-old infants using habituation/dishabituation paradigms with preferential looking methods. In the first phase of the experiment, the author presented objects at specific distances to habituate infants to certain object sizes. In the second phase of the experiment, the author used two objects to test whether the infants looked preferentially to the object with a novel physical or retinal size. One of the objects was physically identical to the habituated object, but it had a different retinal size. Conversely, the second object had the same retinal size as the habituated object but had a different physical size. On average, the infants looked preferentially to the object with a novel physical size. Although this finding indicates that size constancy has an important innate component, studies comparing the degree of size constancy in children with adults revealed that children tend to underestimate objects at a greater distance (Brislin & Leibowitz, 1970; Jenkin & Feallock, 1960; Leibowitz, Pollard, & Dickson, 1967; Zeigler &

Leibowitz, 1957). For example, Zeigler and Leibowitz (1957) asked children and adults to report the perceived size of a stick at various distances and demonstrated that children's responses regressed away from size constancy as viewing distance increased. Therefore, size constancy abilities improve with age, especially for objects placed at far distances.

The developmental profile of perceptual rescaling of stimuli in 2D backgrounds share similarities with the developmental profiles of size constancy (Brislin & Leibowitz, 1970; Leibowitz & Judisch, 1967; Wilcox & Teghtsoonian, 1971). Based on these observations, some have argued that size constancy and the perceptual rescaling of stimuli in 2D backgrounds with pictorial depth cues share similar underlying mechanisms (Leibowitz & Judisch, 1967). For example, Wilcox and Teghtsoonian (1971) compared the degree to which pictorial depth cues influence perceptual size rescaling performance in 3-year-olds, 9-year-olds, and adults. Their results revealed that the presence of pictorial depth cues (e.g., texture gradients and linear perspective cues) affected the perceived size of the stimuli in children aged 9 years and adults but not in children aged 3 years.

In line with this, Leibowitz and Judisch (1967) demonstrated that the degree to which the apparent size of the stimulus at the open end of the pencil-of-lines background (Fig. 5a) increased as a function of age from 3.5 to 13 years, while it remained stable after adolescence. As this developmental pattern was similar to the developmental pattern of size constancy (Zeigler & Leibowitz, 1957), the authors concluded that misapplied size constancy theory explains age-related changes in Ponzo-like illusions. Their results also demonstrated that the strength of the illusion decreased with age from 50 to 88 years. The decline in the strength of the illusion might be explained by age-related declines in the mental representation of 3D information (Plude et al., 1986).

Hadad (2018) has provided support for Leibowitz and Judisch's (1967) study by demonstrating that 4-year-olds and 7-year-olds experience a weaker classic Ponzo illusion than adults ranging in age from 24 to 30 years. Contrary to Leibowitz and Judisch's findings, the author demonstrated that 4-year-olds experienced just as strong an illusion as 7-year-olds. In line with Hadad's finding, there are studies showing no differences among participants ranging in age from 6 years to 14 years (Chevrier & Delorme, 1983; Newman, 1969; Wohlwill, 1962). The discrepancies between these studies and the Leibowitz and Judisch's (1967) study may be due to differences in the methods of these investigations (Pressey, 1987). Namely, differences in the background images that were used to measure the magnitude of the illusion, orientation of the illusion, measurement, sample size, and age range might have led to mixed results.

In another recent study, Cretienoud et al. (2020) examined the effects of pictorial depth cues and age on the magnitude of the Ponzo-like illusions using an

**Table 2** Summary of previous studies examining the effect of age on the magnitude of the Ponzo-like illusions and depth perception

Author(s), year	Age (years)	N	Stimuli	Method	Results
Zeigler and Leibowitz (1957)	7–9, 1–8–24	13	A standard stick placed at various distances between 10 and 100 ft, and a comparison stick placed at 5 ft	Matching method: The task was to match the height of the comparison stick to the height of the standard stick.	Adults but not children demonstrated perfect size constancy. Based on this finding, the authors suggested that size constancy increased with age.
Wohlwill (1962)	7–20	96	The backgrounds with textures that varied in randomness and density	Matching method: Participants were asked to match the size of the comparison rectangle to the size of the standard. The perceived distance was measured by asking participants to report when they saw the moving dot at midway between the screws placed at the near and far ends of the background.	Age affected distance but not size judgments. Specifically, the perceived distance increased with an increase in density and with a decrease in randomness among 7-year-old children. The effect of density and randomness of texture elements on the perceived distance was weaker among adults.
Leibowitz and Judisch (1967)	3.5–88	427	A pencil-of-lines background (Fig. 1e)	Method of constant stimuli: The task was to decide whether the comparison line, presented at the apex of the background image, was smaller or bigger than the standard line, presented at the open end of the background image.	Perceptual size rescaling increased with age from 4 to 13 years, remained stable between 14 and 50 years, and decreased from 50 to 88.
Newman (1969)	6, 10, 14	48	A corridor background (e.g., Fig. 2a)	Matching method: The participants' task was to judge when the bottom line's height appeared equal to the height of the top standard line.	The illusion was present regardless of age; 6-year-olds experienced a strong illusion, but cannot interpret depth.
Wilcox and Teghtsoonian (1971)	3, 9, adults (~2–0)	18	The backgrounds that signal different depths	Implicit positive reinforcement task: In the training phase, two geometrically similar stimuli were presented at different sizes. The participants were encouraged to press the corresponding key when the physically larger stimulus was presented at the left or right side of the display. In the test phase, how the participants respond to the stimulus that was presented at a greater depth was recorded.	In the training phase, both 9-year-olds and adults performed better than 3-year-olds. In the test phase, pictorial depth cues affected perceptual size judgments of 9-year-olds and adults but not 3-year-olds. The authors argued that some pictorial depth cues might affect perceptual size judgments of 3-year-olds, but a larger sample size is necessary to compare the distinct effects of each pictorial depth cue.
Quina and Pollack (1972)	7–19	70	A Ponzo background (Fig. 1a) with contextual lines converged at the display's left side	Method of limits: The standard line was presented at either the open end or apex of the background image. The task was to decide whether the comparison line, presented outside of the Ponzo background, was smaller or bigger than the standard line.	The magnitude of the illusion for the stimulus presented near the apex decreased then increased with an increase in age. The magnitude of the illusion for the stimulus presented at the open end of the background image increased then decreased with age.
Pressey (1974)	5–17	217	The Ponzo (Fig. 1a) and pencil-of-lines (Fig. 1e) backgrounds	Method of limits: The standard line was presented in the upper section of the background image. The task was to decide whether the comparison line, presented outside of the Ponzo background, was smaller or bigger than the standard line.	The magnitude of both the Ponzo and pencil of lines illusions decreased with an increase in age.
Brislin (1974)	3–22	320	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d), and pencil-of-lines (Fig. 1e) backgrounds	Method of constant stimuli: The task was to decide whether the comparison line, presented at the open end of the background image, was smaller or bigger than the standard line, presented at the apex of the background image.	Magnitudes of the field, railroad, and pencil of lines illusions increased with an increase in age. The Ponzo illusion showed an unclear developmental profile.
Wagner (1977)	6–22	384	The Ponzo (Fig. 1a), field (Fig. 1c), and railroad (Fig. 1d) backgrounds	Method of constant stimuli: The task was to decide whether the comparison line, presented at the lower section of the background image, was smaller or bigger than the standard line, presented at the	Magnitudes of the field and railroad illusions increased with an increase in age. The magnitude of the Ponzo illusion decreased with an increase in age.



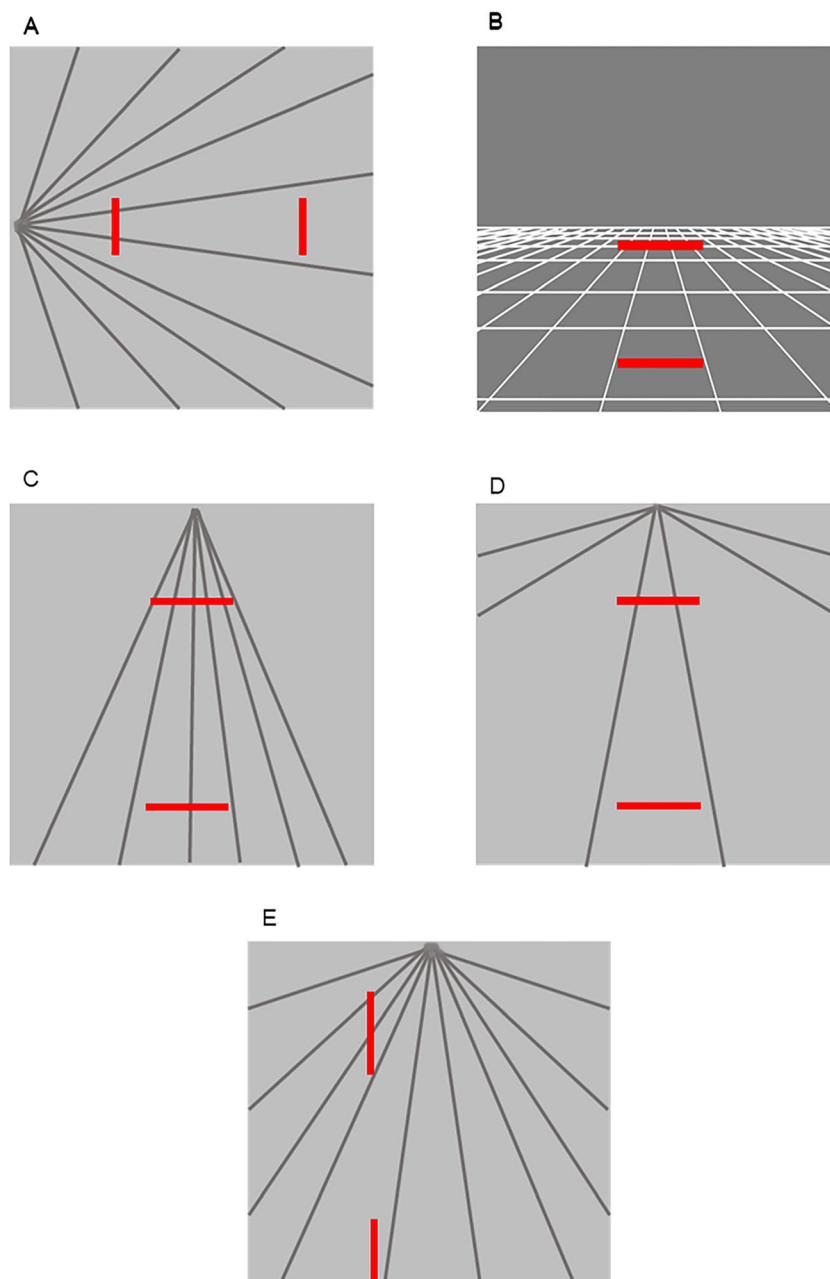
**Table 2** (continued)

Author(s), year	Age (years)	<i>N</i>	Stimuli	Method	Results
Chevrier and Delorme (1983)	6–14	80	The backgrounds with linear perspective, texture gradient, and interposition pictorial depth cues (Fig. 2d)	upper section of the background image. Matching method: The perceived size of the bottom standard triangle was measured by asking participants to match the size of the top comparison triangle to the size of the standard. The perceived distance was measured by asking participants to move the light beside the little girl drawing, placed at the top triangle's position. Both the stimuli and background were presented inside Pandora's box. Flatness depth cues were suppressed by using a transparent background image.	The illusion was present regardless of age. Size illusions increased with an increase in the number of cues. In Pandora's box, perceived depth increased with age and changed depending on the type of cue present. Notably, texture gradients seemed to be important to estimate depth from 2D images with pictorial depth cues. There was a significant relationship between the size and distance adjustments for the older participants (age ranged between 11 and 14).
Predebon (1985)	5–11, 18	212	The Ponzo (Fig. 1a) and pencil-of-lines (Fig. 1e) backgrounds	Method of adjustment: The perceived size of the top standard line was measured by asking participants to match the size of the comparison line that was presented outside of the background image to the size of the standard line.	The magnitude of the Ponzo illusion decreased, then increased with an increase in age. Effect of age on the illusion disappeared when the standard line was presented over the pencil of lines background.
Grzeczowski et al. (2017)	6–81	113	The Ebbinghaus, Müller-Lyer, simultaneous contrast, corridor, White, and tilt illusions	Method of adjustment: In the corridor illusion, the perceived size of the top standard disk was measured by asking participants to match the size of the comparison disk that was presented over background image to the size of the standard disk.	The magnitudes of all illusions decreased with age.
Hadad (2018)	4–5, 7–8, adults (~2–6)	60	The Ponzo (Fig. 1a) background	Method of adjustment: The perceived size of the top and bottom standard lines was measured by asking participants to match the size of the comparison line that was presented outside of the background image to the size of the standard line.	The magnitude of the Ponzo illusion increased with an increase in age; 4-year-olds and 7-year-olds experienced a weaker Ponzo illusion than adults ranging in age from 24 to 30 years; 4-year-olds experienced just as strong an illusion as 7-year-olds.
Cretenoud et al. (2020)	6–66	76	The Ponzo (Fig. 1a), grid (Fig. 5b), and railroad (Fig. 1d) backgrounds	Method of adjustment: The perceived size of the top and bottom standard lines was measured by asking participants to match the size of the comparison line that was presented outside of the background image to the size of the standard line.	The magnitude of the Ponzo illusion for the top stimulus slightly decreased with an increase in age. The effect of the number of pictorial depth cues on the strength of the illusion did not vary as a function of age.

indirect comparison task in which the comparison stimulus was presented outside of the background image. Contrary to Leibowitz and Judisch's (1967) findings, Cretenoud et al. (2020) reported a slight decrease in the magnitude of the railroad illusion with age. In line with this finding, Pressey (1974a) reported that the magnitude of the classic Ponzo and pencil-of-lines illusions for the top stimulus decreased with age when an indirect comparison task was used.

Using a direct comparison task, in which both the standard and comparison stimuli are presented over the background image, Brislin (1974) demonstrated that the strength of the

field and railroad illusions increased with age in Pennsylvanian participants. The author reported that the developmental profile of the classic Ponzo illusion was unclear. In a similar vein, Wagner (1977) demonstrated that the strength of the field and railroad illusions increased with age in Moroccan participants. Additionally, the author showed that the strength of the classic Ponzo illusion decreased with age. Wagner has argued that these findings fit better with mechanisms proposed by Piaget (1969) than with misapplied size constancy theory (Gregory, 1963, 1968, 1998). According to Piaget (1969), visual illusions can be classified into two categories: Type I illusions, which



**Fig. 5** Background images used to test the effect of previous experience on the magnitude of the illusions. The red lines are physically identical. **a** However, we see the left line as being larger than the right line. A similar background was used in the Leibowitz and Judisch (1967) study. Similarly, we see the top line as being larger than the bottom line in **(b)**

**(c–d)** various other pencil-of-lines illusions. **e** The magnitude of the illusion decreases when vertical, rather than horizontal red lines are presented over a pencil-of-lines background. Similar backgrounds were used in the Gandhi et al. (2015) study. (Color figure online)

are supposed to be innate, and Type II illusions, which are thought to be acquired through daily experience with perspective cues. This model predicts that the strength of Type I illusions tends to decrease, whereas the strength of Type II illusions tends to increase with age. Piaget explained developmental difference in Type I and Type II illusions by an increasing number of eye movements. Specially, Piaget suggested that the duration of centration upon stimuli placed over the illusory background declines when scanning strategies

emerge throughout decentration phase (Gardner & Long, 1960). Therefore, according to this model, the decline in the duration of centration increases the magnitude of Type II but not Type I illusions.

Taken together, there are reasons to suggest that the Ponzo illusion might have an important innate component. Supporting and confirming this hypothesis, children and adolescents with dense congenital bilateral cataracts were found to be susceptible to the Ponzo-like illusions when they were

tested a few days after their cataract-removal surgery (see Fig. 5b–e; Gandhi et al., 2015, but also see Fine et al., 2003; Lazar, 1964). Because these children have poor spatial vision (Andres et al., 2017), it might be argued that a limited visual experience is sufficient for being susceptible to these illusions. This finding is in line with studies showing no impact of cultural differences in the strength of the classic Ponzo illusion.

To sum up, there is still no agreement on the developmental trajectory of the Ponzo-like illusions. Although some of the cross-cultural studies reviewed above would predict that young children might experience the classic Ponzo illusion and that its strength would be unaffected by the exposure to the different pictorial depth cues, not all studies reported in this section are in agreement with this prediction. The evidence is mixed on if and how age affects Ponzo-like illusions.

### Alternative theories that explain the Ponzo illusion with mechanisms unrelated to depth perception

As one can see from Tables 1 and 2, studies examining the effect of culture and age on the magnitude of the Ponzo-like illusions and depth perception have yielded mixed results. Many have speculated that either low-level assimilation-contrast effects and eye movements, or high-level perceptual mechanisms that help us to perceive objects as having the same physical features when they are viewed from different angles, hereafter referred to as tilt constancy, might explain the classic Ponzo illusion better than misapplied size constancy theory. Empirical arguments for and against the contour-proximity (Fisher, 1968a, 1969, 1970, 1973), pool-and-store (Girgus & Coren, 1982), assimilation (Pressey, 1974b; Pressey et al., 1971; Pressey & Epp, 1992), and tilt constancy (Prinzmetal & Beck, 2001; Prinzmetal et al., 2001) theories of the classic Ponzo illusion are reviewed in this section. Table 3 provides a summary of alternative theories that explain the classic Ponzo illusion with mechanisms unrelated to depth perception. Note that the proposed alternative explanations are not mutually exclusive.

### Contour proximity theory

The contour proximity theory, originally formulated by Fisher (1968a, 1969, 1970, 1973), asserts that the classic Ponzo illusion is generated by the distances between the endpoints of stimuli and the sides of the contextual converging lines, such that longer distances between the two decrease the apparent size of the stimuli. To provide evidence for his theory, Fisher (1968c) tested participants with a background image which could be judged either as a corridor image in which the

stimulus near the apex appears farther away than the stimulus near the base of the trapezoid-like shape or as a pyramid image in which the stimulus near the base of the trapezoid-like shape appears farther away than the stimulus near the apex (see Fig. 6a). The author showed that the strength of the illusion was similar between the participants who were instructed to perceive the background image as a pyramid and those who were instructed to perceive the background image as a corridor (see also Fisher, 1970). These findings suggest that it is the physical distance between contextual converging lines and the endpoints of the stimuli, rather than the different perceived distance of the stimuli, that drives the illusion, a conclusion that is at odds with misapplied size constancy theory.

Fisher (1968a, 1969, 1970, 1973) has provided further evidence for his theory by showing how the angle of the contextual converging lines and the distance between the standard and comparison lines can affect the magnitude of the classic Ponzo illusion. For example, the author reported that the strength of the illusion was larger when the endpoints of the stimulus were closer to the sides of the contextual converging lines (with contextual lines converging at 45 degrees) than when they were farther away (with contextual lines converging at 105 degrees; Fisher, 1969). According to the contour proximity theory, the progressive increase in the perceived size of the stimulus from the bottom to the top positions of the classic Ponzo background results from an increase in proximity to the contextual converging lines (Fisher, 1968b, 1969).

Additional support to the contour proximity theory comes from studies that have shown how the strength of the classic Ponzo illusion varies as a function of the orientation of the stimuli, which are presented over the background (Gilliam, 1973; Schiffman & Thompson, 1978). For example, Schiffman and Thompson (1978) showed that the classic Ponzo illusion with contextual lines converging in the upper visual field occurs only when horizontal stimuli were presented over the background (Fig. 1a). This finding was in line with Fisher's theory, predicting that converging contextual lines would affect the perceived length of the horizontal (Fig. 1a) but not the perceived height of the vertical (Fig. 6b) stimuli in these instances.

Perceived depth changes as a function of the angle of the contextual converging lines and the distance between the standard and comparison lines. Therefore, the findings showing how these two variables affect the magnitude of the classic Ponzo illusion do not contradict Gregory's misapplied size constancy theory. Moreover, Gregory outlined a typical view hypothesis that might explain why the background image in Fig. 6a produced the same amount of the illusion for the apparently near and far stimuli even when the background image was perceived as a pyramid rather than a corridor (for a review, see Green, 1972). Namely, Gregory suggested that primary perceptual size rescaling mechanisms underly



**Table 3** Summary of alternative theories that explain the classic Ponzo illusion with mechanisms unrelated to depth perception

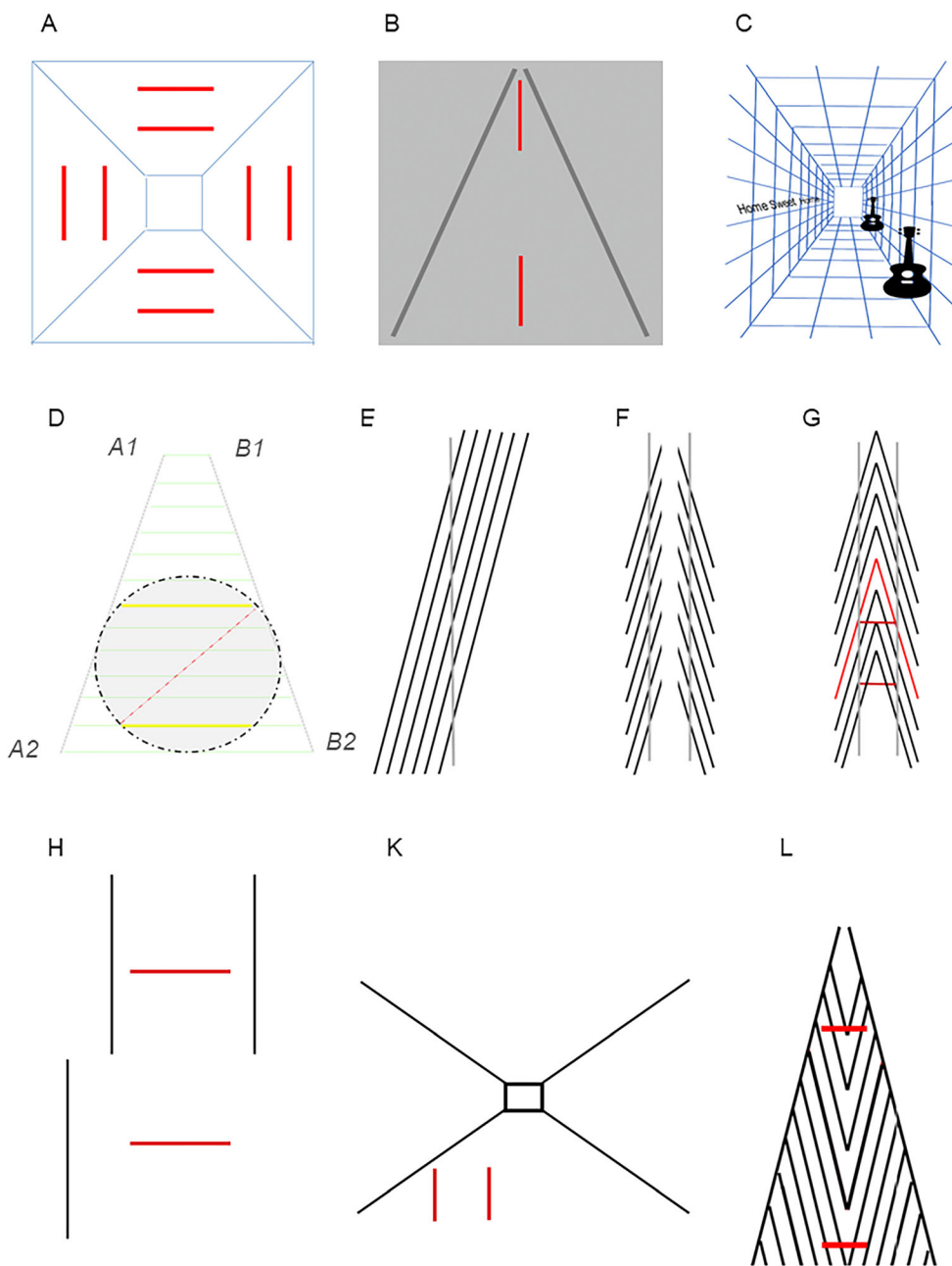
Theory	Theoretical explanation	Empirical evidence supporting the theory	Empirical evidence against the theory
Contour proximity theory	Differences in the distances between the endpoints of stimuli and the sides of the contextual converging lines drive the classic Ponzo illusion. Specifically, longer distances between the two decrease the apparent size of the stimuli.	The strength of the illusion was similar between the participants who were instructed to perceive the background image in Fig. 6a as a pyramid and those who were instructed to perceive the same image as a corridor (Fisher, 1968c). Moreover, the strength of the classic Ponzo illusion varies as a function of the orientation of the stimuli, which are presented over the background (Gilliam, 1973; Schiffman & Thompson, 1978).	The classic Ponzo illusion was observed under conditions of sequential presentation (Shen et al., 2015). Moreover, the classic Ponzo background produced a strong illusion when the background and stimuli were presented separately to the left and right eyes (Song et al., 2011; Yildiz et al., 2021b).
Pool-and-store theory	In a classic Ponzo background, the bottom but not the top stimulus requires successive eye movements to be processed. Eye movements consequently cause the stimulus to appear smaller.	The strength of the contrast but not the assimilation illusion was larger when more time is given to inspect the illusion (Girgus & Coren, 1982).	The classic Ponzo illusion was observed under conditions of sequential presentation (Shen et al., 2015). A study using negative afterimages has demonstrated that the classic Ponzo illusion was experienced when both the stimulus and background were fixed to the retina (Qian et al., 2016).
Assimilation theory	In the classic Ponzo illusion, perceived size of the stimulus assimilates toward the contextual horizontal magnitudes while the contextual horizontal magnitude assimilates toward the mean. The effect of the contextual magnitudes on the perceived size of the stimulus decreased when the contextual lines did not fall within the attentive field.	The perceived size of the top stimulus was underestimated when the stimulus was presented over contextual lines converging at 150 degrees.	The lightness contrast between the contextual converging lines and the standard stimulus affected the strength of the classic Ponzo illusion (Jaeger et al., 1980).
Tilt constancy theory	The classic Ponzo illusion relies on tilt constancy mechanisms while the corridor illusion relies on size constancy mechanisms. The tilt-induction effect occurs at the left and right endpoints of the top and bottom stimuli in the classic Ponzo illusion.	The strength of the classic Ponzo, but not the corridor illusion increased when observers' body postures were tilted 30 degrees counter-clockwise (Prinzmetal & Beck, 2001). The vertical line that was presented near the sidewall perceived larger than the other (Prinzmetal et al., 2001).	In Fig. 6j, the illusory effect disappears when the top and bottom lines are presented over the classic Ponzo background with oblique lines that converge in the opposite direction (Roncato et al., 1998). The theory cannot explain why the size of the top but not the bottom stimulus tends to be perceived differently than its physical size when the comparison stimulus is presented outside of the Ponzo background (Cretienoud et al., 2020; Yildiz et al., 2019, 2021a).

illusory size perception in the Ponzo-like illusions and that these mechanisms work regarding familiarity principles rather than cognitive decision-making processes. In other words, even if we know that the lines are physically identical when they are presented over the pyramid background, our visual system processes the background as a corridor rather than a pyramid, and perceptually rescales the size of the apparently far and near objects so that we perceive the latter as having a smaller size than the former. Nevertheless, in the misapplied size constancy theory, Gregory predicted to observe an equal amount of illusion for lengths of the horizontal (Fig. 1a) and heights of the vertical (Fig. 6b) stimuli. Therefore, the Schiffman and Thompson's (1978) findings argue against misapplied size constancy theory. In line with the predictions

of misapplied size constancy theory, Prinzmetal and Beck (2001) showed that the height of the stimuli presented at the far end of the corridor background in Fig. 6c was perceived as larger than its physical height. Why the classic Ponzo and corridor backgrounds affect perceived height of the vertical stimuli differently remains unknown.

### Pool-and-store theory

Fisher's contour proximity theory inspired other theories on the classic Ponzo illusion. For example, the pool-and-store theory proposed by Girgus and Coren (1982) furthered Fisher's theory by explaining the underlying mechanisms of



**Fig. 6** Background images used in studies that explained the Ponzo illusion with mechanisms unrelated to depth perception. **a** A background image which can be judged either as a corridor or a pyramid image. A similar variant was used in the Fisher (1968c) study. **b** A Ponzo background with the two vertical lines that have the same retinal image size. A similar variant was used in the Schiffman and Thompson (1978) study. **c** A corridor background with linear perspective and relative size pictorial depth cues. A similar variant was used in the Prinzmetal and Beck (2001) study. **d** A Ponzo illusion with contextual horizontal magnitudes and attentive field. The two yellow lines have the same retinal image size. Yet we see the upper line as being larger than the lower line. In the assimilation theory, Pressey (1974b) explained the illusory effect with contextual horizontal magnitudes (green horizontal lines) and the attentive field (circular area fall within the dashed circle). Dashed red line demonstrates the diameter of

the minimum attentive field. **e**Tilt-induction illusion. A vertical line appears as tilted in the counterclockwise direction when it was presented over contextual lines which are tilted in the clockwise direction. **f** Zöllner illusion. The left vertical line appears as tilted in the counterclockwise direction when it is presented over contextual lines which are tilted in the clockwise direction. In contrast, the right vertical line appears as tilted in the clockwise direction when it is presented over contextual lines which are tilted in the counterclockwise direction. **g** A Ponzo illusion which is hidden in the Zöllner illusion. **h** Rectilinear Ponzo illusion. A similar background was used in the Prinzmetal et al. (2001) study. **i** A hallway with linear perspective cues. A similar background was used in the Prinzmetal et al. (2001) study. **j** A Ponzo background with the oblique lines converged in the opposite direction. A similar background was used in the Roncato et al. (1998) study. (Color figure online)

the contour proximity theory. According to Girgus and Coren, eye-movements can account for the size distortions: When the gap between the endpoints of the stimulus and the contextual converging lines is small, both the stimulus and the background information fall on the retina at the same time and are registered together, whereas when the gap between the endpoints of the stimulus and contextual converging lines is large, then the stimulus and the background cannot be registered together without successive eye movements. Thus, in a classic Ponzo background, the bottom but not the top stimulus requires successive eye movements to be processed. Eye movements consequently cause the stimulus to appear smaller.

In agreement with the contour proximity theory, the pool-and-store theory states that the top stimulus appears larger than its physical size because the endpoints of the top stimulus are close to the sides of the contextual converging lines. This expansion of the top stimulus is regarded as an assimilation illusion. In contrast, the bottom stimulus appears smaller than its physical size because the endpoints of the bottom stimulus are further away from the sides of the contextual converging lines. This shrinkage of the bottom stimulus is regarded as a contrast illusion.

Since Girgus and Coren (1982) explained the assimilation illusion with visual information received in a single glance, the pool-and-store theory predicts that when the classic Ponzo background and the stimuli are presented sequentially to the viewers, a contrast but not an assimilation illusion should be observed. Contrary to this hypothesis, a recent study has reported an assimilation illusion in the classic Ponzo background under conditions of sequential presentation (Shen et al., 2015). Moreover, another study using negative afterimages has demonstrated that the classic Ponzo illusion was experienced even when both the stimulus and background were fixed to the retina (Qian et al., 2016). The observations of size distortions during sequential presentations and retinal afterimages argue against the pool-and-store theory. Finally, it should be noted that the pool-and-store theory was envisaged to explain the classic version of the Ponzo illusion. Therefore, it is unclear whether it can be extended to other versions of the illusion that do not include converging lines, such as the field illusion, or whether it can explain why the magnitude of the illusion increases with an increase in the number of pictorial depth cues.

## Assimilation theory

Assimilation theory, originally formulated by Pressey and his colleagues (1974b, 2013, 1971) asserts that whenever the participant judges the perceived size of a stimulus presented over the classic Ponzo background, the brain assigns a horizontal magnitude to each point in between the contextual converging lines (the so-called contextual magnitude, see green horizontal

lines in Fig. 6d). The theory is based on three assumptions (Pressey, 1974b, 2013; Pressey et al., 1971). First, the perceived size of the stimulus assimilates toward the contextual horizontal magnitudes while the contextual horizontal magnitude assimilates toward the mean. Namely, extremely small contextual horizontal magnitudes (horizontal magnitude between A1 and B1 in Fig. 6d) are overestimated and cause an apparent expansion of the top standard stimulus, while extremely large magnitudes (horizontal magnitude between A2 and B2 in Fig. 6d) are underestimated and cause an apparent shrinkage of the bottom standard stimulus. Second, contextual horizontal magnitudes have stronger effects on the perceived size of the stimulus if they fall within the “attentive field.” In assimilation theory, the attentive field is hypothesized as a circular area in which the visual system integrates temporally and spatially separable samples of visual information by taking their weights into account. The diameter of the minimum attentive field is initially defined as the distance between the two extreme ends of the standard and comparison stimuli (dashed red line in Fig. 6d; Pressey, 1974b; Pressey et al., 1971). Third, the magnitude of the classic Ponzo illusion for the top and bottom stimuli increases with an increase in the range of the contextual horizontal magnitudes that fall within the attentive field (Pressey, 1972). In other words, the magnitude of the illusion changes as a function of the distance between extremely small and large magnitudes that fall within the attentive field.

According to assimilation theory, if the task requires participants to decide whether the perceived size of the bottom comparison stimulus in Fig. 6d is smaller or larger than the top standard stimulus, then the participants would repeatedly attend to the top and then the bottom stimuli. As the contextual horizontal magnitudes become larger from the top to the bottom stimulus, the size of the top standard stimulus is overestimated. Conversely, if the task requires participants to decide whether the perceived size of the top comparison stimulus is larger or smaller than the bottom standard stimulus, then the participants would repeatedly attend to the bottom and then the top stimuli. As the contextual horizontal magnitudes become smaller from the bottom to the top stimulus, the size of the bottom stimulus is underestimated.

Studies showing that the strength of the illusion changes as a function of the angle between converging contextual lines (Pressey et al., 1971) and where the comparison stimulus was presented (Pressey, 1974b) provide evidence for assimilation theory. For example, Pressey et al. (1971) revealed that the perceived size of the top stimulus was underestimated when the stimulus was presented over contextual lines converging at 150 degrees. This finding suggests that the effect of the contextual magnitudes on the perceived size of the stimulus decreased when the contextual lines did not fall within the attentive field.

Although Pressey explained all size distortions in the classic Ponzo background with assimilation-related mechanisms, the differences in the relationship between age and the magnitude of the classic Ponzo illusion for the top and bottom stimuli suggested to the author that the mechanisms underlying the apparent expansion and shrinkage of the stimuli might be different (Pressey, 1974a). In fact, as discussed above (see section *The effects of age on the magnitude of the illusion*), the strength of the classic Ponzo illusion tends to increase or decrease during development supposedly due to differences in task demands. In the case of Pressey's study (1974a), the perceived size of the top stimulus in the classic Ponzo illusion decreased with development. This finding is in line with studies demonstrating how the strength of assimilation illusions decreases with age (Cretenoud et al., 2020; also see Predebon, 1985; Quina & Pollack, 1972). Pressey argued that Leibowitz and Judisch's (1967) results might be a replication of studies demonstrating how the strength of contrast illusions increases with age.

Jaeger et al. (1980) provided evidence against Pressey's assimilation theory by showing that the length of the contextual converging lines and the lightness contrast between the contextual converging lines and the standard stimulus can affect the strength of the classic Ponzo illusion. Specifically, the authors showed that if both the contextual converging lines and the stimulus were presented in grey, then shorter contextual converging lines produced a stronger illusion than the longer contextual converging lines for the stimulus near the vertex. As the strength of the illusion varied as a function of lightness contrast and wedge length, Jaeger et al. (1980) concluded that low-level contour interactions might play an important role in the classic Ponzo illusion. Such a conclusion contradicts Pressey's assimilation (Pressey, 1972; Pressey et al., 1971, 1974b) and Gregory's misapplied size constancy (1963, 1968, 1998) theories, while it is more in line with Fisher's contour proximity theory (1968a, 1969, 1970).

Contrary to Jaeger et al.'s (1980) interpretation, recent evidence suggests that nonlinear increases in the magnitude of the classic Ponzo illusion occur with an increase in inducer contrast as a result of high-level visual processing through extrastriate cortical areas, rather than low-level visual processing through the striate cortex (Brown et al., 2018). Therefore, both assimilation and misapplied size constancy theories might account for these findings based on high-level visual processing mechanisms.

## Tilt constancy theory

Prinzmetal and Beck (2001) proposed an alternative theory for the classic Ponzo illusion that relies on tilt constancy mechanisms. Tilt constancy mechanisms operate so that we perceive vertical lines as vertical even when their retinal orientations

change as a function of our body posture while we are tilting our heads. Tilt constancy theory has its roots in the tilt-induction effect (Gibson & Radner, 1937), where a vertical line appears tilted in the counterclockwise direction when it is presented over contextual lines that are tilted in the clockwise direction (Fig. 6e). Thus, the tilted context distorts participants' orientation judgments.

Tilt constancy theory can explain a number of perceptual phenomena, such as tilt-induction effect, the classic Ponzo, Zöllner, Pogendorff, Wündth-Hering, and café-wall illusions. For example, in Fig. 6f, the left vertical line appears tilted in the counterclockwise direction when it is presented over contextual lines which are tilted in the clockwise direction. In contrast, the right vertical line appears tilted in the clockwise direction when it is presented over contextual lines which are tilted in the counterclockwise direction. This illusion is known as the Zöllner illusion. When the gaps between the contextual lines are filled, different converging lines are obtained (Fig. 6g). The red lines in Fig. 6g illustrate a classic Ponzo illusion. The absence of vertical lines in the classic Ponzo illusion (Fig. 1a) casts doubt on the generality of a tilt constancy explanation. Prinzmetal and colleagues suggested that the tilt-induction effect occurs at the left and right endpoints of the top and bottom stimuli in the classic Ponzo illusion. Although the authors argued that processing arising from cortical lateral inhibitions in early visual areas of the cortex (Blakemore et al., 1970) could not explain tilt constancy, how the visual system integrates the endpoints of the top and bottom stimuli to enable tilt constancy remains unclear.

Support for the tilt constancy theory comes from a study by Prinzmetal and Beck (2001) showing that the strength of the classic Ponzo, but not the corridor illusion increased when the observers' body postures were tilted 30 degrees counterclockwise. These findings indicated that there must be different mechanisms underlying the classic Ponzo and the corridor illusions. In particular, the authors proposed that size constancy mechanisms might drive the corridor illusion, while tilt constancy mechanisms might underlie the classic Ponzo illusion. The increase in the strength of the Ponzo illusion at 30 degrees rotation seems to suggest that feedforward projections in the visual system play a greater role than extravisual cues in size judgments when the participants were tilted (e.g., vestibular and somatosensory information). Therefore, in disagreement with misapplied size constancy theory, which explains the size distortions in the classic Ponzo illusion with feedback projections (Sperandio & Chouinard, 2015), tilt constancy theory explains size distortions by feedforward projections.

To compare the predictions of tilt constancy theory with Girgus and Coren's (1982) pool-and-store theory, Prinzmetal et al. (2001) presented the top stimulus over a smaller rectangular area using a rectilinear background image (Fig. 6h). The authors reasoned that if the illusory size perception in the classic Ponzo illusion is related to the presence of oblique



contextual lines, then the rectilinear background would produce little or no illusion. Contrarily, if the illusory size perception in the classic Ponzo illusion is related to the difference in the distances between the contextual lines and the stimuli, then the magnitude of the illusion in the rectilinear background would be equal to the magnitude of the illusion in the classic Ponzo background. Their results provided evidence for the tilt constancy theory by supporting the first possibility.

To compare the predictions of the tilt constancy theory with the misapplied size constancy theory, the same authors presented the two vertical lines over a Ponzo-variant background with contextual lines (Fig. 6i). Although the authors claimed to the contrary, their background seems to us as a corridor drawing. The authors reasoned that if the illusory size perception in the Ponzo-variant illusion is related to the presence of oblique contextual lines, then the vertical line that was presented near the sidewall would be perceived larger than the other. Contrarily, if the illusory size perception in the Ponzo-variant illusion is related to the presence of linear perspective depth cues, then the vertical line that was presented near the sidewall would be perceived as having the same size as the other line. Their results provided evidence for the tilt constancy theory by supporting the first possibility. Indeed, this is an interesting finding that needs to be replicated using different Ponzo-like illusions. As a reminder from the authors' earlier discussion (Prinzmetal & Beck, 2001), the tilt constancy theory predicts that the corridor but not the classic Ponzo illusion is driven by size constancy mechanisms. Why the authors chose a background that can easily be interpreted as a corridor drawing to compare their theory with misapplied size constancy theory remains unknown.

Prinzmetal et al. (2001) has shown that the orientation of the “virtual” line between the endpoints of the top and bottom stimuli was affected by the nearest contextual converging lines. Therefore, according to tilt constancy theory, if the oblique lines within the classic Ponzo background converge at the opposite direction as illustrated in Fig. 6j, then the size of the bottom stimulus should be overestimated, while the size of the top stimulus would be underestimated. Contrary to the predictions of the tilt constancy theory, Roncato et al. (1998) demonstrated that the illusory effect disappears when the top and bottom lines are presented over the classic Ponzo background with oblique lines that converge in the opposite direction. Moreover, in a recent study, Cretienoud et al. (2019) demonstrated that there were strong within-illusion correlations when the classic Ponzo illusion was tested under different orientations. This finding contradicts with the Prinzmetal and Beck (2001) finding demonstrating that the classic Ponzo, but not the corridor illusion increased when the participants' body postures were tilted 30 degrees counterclockwise. Finally, the “virtual line” assumption in tilt constancy theory cannot explain why the size of the top but not the bottom stimulus tends to be perceived differently than its physical size

when the comparison stimulus is presented outside of the Ponzo background (Cretienoud et al., 2020; Yildiz et al., 2019, 2021a).

### Are the alternative theories able to explain the inconsistent findings?

None of the alternative theories presented in this section has predictions about if and how the number of pictorial depth cues affects the illusion's magnitude. Although inconsistent findings on the Ponzo-variant illusions cannot be explained by these alternative theories, they might account for some of the discrepancies on the classic Ponzo illusion. For example, in line with these alternative theories, the literature reviewed demonstrated that the strength of the classic Ponzo illusion did not change depending on previous experience with pictorial depth cues (Brislin, 1974; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Segall et al., 1966; Wagner, 1977). Therefore, low-level mechanisms such as lateral inhibition or high-level mechanisms that are unrelated to depth perception might be at play in the classic Ponzo illusion. Since converging evidence has demonstrated that the classic Ponzo illusion requires the involvement of binocular neural populations in the primary visual cortex and higher visual areas (Chen et al., 2018; Song et al., 2011; Yildiz et al., 2021b), it is hard to explain the classic Ponzo illusion with solely low-level mechanisms as contour proximity theory would claim (Fisher, 1968a, 1969, 1970, 1973). Also, the pool-and-store theory of the classic Ponzo illusion (Girgus & Coren, 1982) has been challenged by recent studies showing an illusion when the afterimages of the classic Ponzo background were fixed to the retina (Qian et al., 2016) or when the background and stimuli were presented sequentially to participants (Shen et al., 2015). Nevertheless, the assimilation (Pressey, 1974b; Pressey et al., 1971; Pressey & Epp, 1992) and tilt-constancy (Prinzmetal & Beck, 2001; Prinzmetal et al., 2001) theories may still contribute to the explanation of the classic Ponzo illusion.

Both the assimilation (Pressey, 1974b; Pressey et al., 1971; Pressey & Epp, 1992) and tilt constancy (Prinzmetal & Beck, 2001; Prinzmetal et al., 2001) theories support the assumption that the underlying mechanisms of the classic Ponzo illusion and its rich-context variants (i.e., the field, railroad, and corridor illusions) are entirely different. This hypothesis was confirmed by one of our previous studies (Yildiz et al., 2021b), where we have shown that the underlying mechanisms of the classic Ponzo and corridor illusion with linear perspective cues are not entirely the same.

In the tilt constancy theory, Prinzmetal and Beck (2001) have explained the size distortions with feed-forward connections while in the assimilation theory, Pressey (2013) has argued that assimilation effects depend on attention. Therefore,

neither theory can be explained by solely low-level mechanisms that occur at an early stage in the processing hierarchy. Since both the assimilation and tilt-constancy theories require the involvement of higher-level cortical areas, the finding showing that the binocular neural populations play a greater role in the classic Ponzo background does not contradict predictions of these two alternative theories (Song et al., 2011; Yildiz et al., 2021b). Similarly, the finding demonstrating that priming effects induced by a Ponzo illusion increase with slower responses does not contradict predictions of these two alternative theories (Schmidt & Haberkamp, 2016).

Taken together, we speculate that the mechanisms outlined in the assimilation (Pressey, 1974b; Pressey et al., 1971; Pressey & Epp, 1992) and tilt constancy (Prinzmetal & Beck, 2001; Prinzmetal et al., 2001) theories might drive the classic Ponzo but not the Ponzo-variant illusions. Moreover, these alternative theories cannot explain culture- and age-related changes in the magnitude of the Ponzo-variant illusions. In the subsequent section, we argue that inconsistent findings on the Ponzo-variant illusions require a reformulation of the misapplied size constancy theory.

## A Bayesian-motivated reconceptualization of misapplied size constancy theory

Evidence has supported the notion that the magnitude of the Ponzo-like illusions changes depending on the number of pictorial depth cues (Brislin, 1974; Fineman, 1981; Leibowitz et al., 1969; Wagner, 1977; Yildiz et al., 2019, 2021a, 2021b) as well as previous experience (Brislin, 1974; Leibowitz et al., 1969; Leibowitz & Pick, 1972; Wagner, 1977). Yet this is not always the case as repeatedly shown in the literature (Segall et al., 1966; Wagner, 1977; Yildiz et al., 2019, 2021a, 2021b). Given these inconsistencies in the literature, one may justifiably wonder why the magnitude of the illusion does not always increase with the number of pictorial depth cues and repeated exposure to the pictorial depth cues, as one would predict if misapplied size constancy theory were true.

In this section, we argue that these discrepant findings require a reformulation of the misapplied size constancy theory. We propose a Bayesian-motivated reconceptualization of the theory that explains all Ponzo-like illusions in terms of prior information and prediction errors. Bayesian approaches have been successful at explaining flash-lag (Khoei et al., 2017) and motion (Weiss et al., 2002) illusions. As emphasized before, Gregory's theory has its roots in Bayesian concepts of prediction errors and prior information (Gregory, 2005, 2006a, 2006b). Yet, to the best of our knowledge, the findings reported in this review have never been discussed under a Bayesian framework.

## Could a Bayesian-motivated reconceptualization of the misapplied size constancy theory explain the Ponzo-variant illusions?

Perceived depth helps us interact with 3D objects in the 3D world. As the visual information that we receive through our eyes is 2D, the perceived depth depends mostly on binocular (e.g., binocular disparity, convergence) and pictorial (e.g., linear perspective cues and textures) sources of depth cues. How the visual system is able to estimate depth from pictorial depth cues is not completely known. Under a Bayesian framework, we may explain the process with prior probability distributions and likelihood functions.

In a Bayesian framework, prior probability distribution or simply a “prior” refers to the probability assigned for different physical states of the world before receiving sensory information. In contrast, the likelihood function refers to the probability of receiving a specific retinal image in a given situation. The normalized product of the prior information with the incoming sensory input gives the posterior probability function (Howe et al., 2006).

Where priors come from is one of the outstanding questions in perception science. Many have agreed that there are innate mechanisms that store some of the essential knowledge in our genetic code (for empiricists, see Chater et al., 2015; for nativists, see Simpson et al., 2005). This innate knowledge, if it exists, might affect how the brain learns environmental statistics to estimate depth using pictorial depth cues (Gregory, 2009). Accumulating evidence suggests that we could learn to adapt multiple priors or modify our existing priors based on experience (Adams et al., 2004; Seydell et al., 2011). Since different priors might be useful in different situations, the visual system does not pick a prior and drop others. Instead, it assigns different weights to different priors based on previous experience (Knill, 2003; Seydell et al., 2011). These weights are used to form a prior probability density function (Seydell et al., 2011).<sup>2</sup>

There is reason to assume that we have innate priors about how two parallel lines recede into the distance and appear as converging lines. Yet, do we have to apply this prior information to estimate depth even when we truly encounter converging lines? Since the visual system could learn to adapt multiple priors or modify our existing ones based on experience, it might also learn to form a prior probability density function that considers the weights of different interpretations of converging lines (i.e., converging lines or parallel lines that converge in the distance as the linear perspective depth cues). A prior probability density function could be formed in a similar vein for the textures, assuming that they might be interpreted in two different ways: uniform or non-uniform texture

<sup>2</sup> The probability density function describes the relative probability that a random variable corresponds to a certain value.

patterns. Together, these priors help us make inferences about the physical distance in the 3D world from 2D retinal images of pictorial depth cues.

We may also have prior information about how two physically identical objects subtend different visual angles on the retina when they are placed at two different distances. Namely, the far object subtends smaller visual angles while the near one subtends larger visual angles. This prior information almost always helps us perceive objects with similar physical size as having the same size despite differences in retinal sizes that are caused by changes in viewing conditions. In this way, we see the world around us as a stable and coherent view. In fact, to maintain size constancy, the brain continuously rescales the object's size with the perceived distance using the prior information about how two physically identical objects subtend different visual angles on the retina when they are placed at different distances. Yet, when the brain uses this prior information to rescale the size of 2D stimuli presented over a background with pictorial depth cues, we experience size illusions.

For example, some of the Ponzo-like backgrounds contain pictorial depth cues that simulate greater depth in the upper section. When the depth information is extracted from these 2D flat images, the brain expects to process smaller and larger retinal sizes for similar objects placed in the upper and lower sections of these backgrounds, respectively. Having physically identical 2D objects in the upper and lower sections of a 2D image with pictorial depth cues violates what would be predicted from physically identical 3D objects placed at different distances. Violation of expectations causes prediction errors. To minimize prediction errors, perceptual rescaling mechanisms operate so that sizes of the objects are rescaled based on their perceived depth. This *inappropriate* perceptual rescaling is useful to minimize prediction errors. However, it causes the objects in the lower and upper sections to appear smaller and bigger, respectively, on Ponzo-like illusory backgrounds displayed on a flat surface.

### **How does misapplied size constancy theory explain the culture-related and age-related changes in the Ponzo-variant illusions with prior information and prediction errors?**

If prior information and prediction errors are indeed important for the Ponzo-like illusions, then one might expect to find increases in the strength of the illusions with an increase in the number of pictorial depth cues. Under a Bayesian framework, the reason for the increase in the magnitude of the illusion with an increase in the number of depth cues is that information from a 2D display of the more abstract version of the Ponzo illusion, where there are less depth cues, could be interpreted in more than one way. In contrast, 2D displays of Ponzo-like illusions with additional pictorial depth cues strongly suggest one interpretation. Indeed, it is conceivable

that the relative certainty of the prior information suggesting that converging lines of the Ponzo background represent linear lines receding in depth is relatively weaker than the relative certainty of the prior information suggesting that converging lines of a railroad represent linear lines receding in depth. In line with this interpretation, previous results suggest that the strength of the illusion increases with an increase in the number of pictorial depth cues only when the available pictorial depth cues signal more reliable depth information (Yildiz et al., 2019, 2021a, 2021b). Therefore, differences in the relative certainty of the prior information affect the magnitude of the illusion.

A Bayesian-motivated reconceptualization of the misapplied size constancy theory could also explain the differences in the interocular transfer effects of the textures and linear perspective cues as being due to differences in the relative certainty of the prior information (Yildiz et al., 2021b). As mentioned earlier, the background with textures might be interpreted in two different ways: as being composed of uniform or nonuniform texture patterns. Similarly, the Ponzo background might be interpreted in two different ways: as simply consisting of converging lines or as consisting of parallel lines that converge in the distance (thus serving as linear perspective cues). Perhaps because there is greater uncertainty, extraction of depth information requires the involvement of higher-order cortical areas in the Ponzo illusion and the corridor illusion with textures. As the background with linear perspective depth cues in the corridor illusion strongly suggests one interpretation over the other, extraction of the depth information clearly does not require the involvement of higher-order cortical areas. Thus, we can state that monocular neural populations play a more important role in the corridor illusion with linear perspective cues.

Indeed, if prior information and prediction errors are important for the Ponzo-like illusions, then one might also expect to find differences in the strength of the illusion depending on the previous experience with pictorial depth cues in the surrounding environment and depending on the previous experience with the 2D representations of the conventional 3D pictorial depth cues used in pictures. Many studies showed that cultural differences affect how the strength of the illusion increases with an increase in the number of pictorial depth cues. As we proposed earlier, the increase in the strength of the illusion with an increase in the number of pictorial depth cues could be related to the increase in the relative certainty of the prior information. Namely, the participants who live in environments with city-blocks, rectangular buildings, and street patterns experience stronger railroad and field illusions, perhaps because these participants have stronger priors about how the sizes of 3D objects change in the distance over fields and railroads (Brislin, 1974; Leibowitz et al., 1969; Wagner, 1977). Conversely, for the participants who have a weaker or no prior information about how specific pictorial depth cues

affect the size of objects in real life, the magnitude of the illusion does not change depending on the number of available pictorial depth cues.

Additionally, cross-cultural studies reveal evidence that even the observers, who have weaker prior information about how linear perspective depth cues affect the size of the objects in real life, are susceptible to the classic Ponzo illusion. Because having a default system that extracts depth information from linear perspective cues is vital to survival, evolutionarily programmed mechanisms might explain this finding. Therefore, innate knowledge about how linear lines recede into depth might affect how the brain learns environmental statistics to estimate depth using pictorial cues. Alternatively, mechanisms outlined in the assimilation and tilt-constancy theories of the classic Ponzo illusion might help the brain learn how to rescale size of a stimulus using contextual converging lines.

In misapplied size constancy theory, Gregory (1963, 1968) proposes that the stimulus placed at a location where the pictorial depth cues signal closer distance appears smaller than its physical size while the stimulus placed at a location where the pictorial depth cues signal greater depth appears larger than its physical size. Yet, studies revealed that the stimulus placed at a location where the pictorial depth cues signal closer distance does not appear smaller than its physical size when the comparison stimulus was presented outside the background image (Cretenoud et al., 2020; Yildiz et al., 2019, 2021a). In our studies, we recorded participants' eye movements and found that participants directed their gaze more frequently to the comparison ring while they were judging the bottom standard ring's size (Yildiz et al., 2019, 2021a). Therefore, eye movements can explain why the stimulus placed at a location where the pictorial depth cues signal closer distance does not appear smaller than its physical size when the comparison stimulus was presented outside the background image.

A Bayesian-motivated reconceptualization of misapplied size constancy theory could explain these inconsistent findings by reference to prior information. We argue that prior information about how two physically identical objects subtend different visual angles on the retina when they are placed at two different distances could explain why the visual system perceptually rescales the size of both the top and bottom rings when the comparison stimulus is presented inside the background image. There is reason to assume that the visual system uses this prior information for the top stimulus even when the comparison is presented outside of the background image. Therefore, a Bayesian-motivated reconceptualization of misapplied size constancy theory could explain why the size of the top stimulus changes depending on the presence of pictorial depth cues by reference to prior information (Yildiz et al., 2019, 2021a). Previous findings suggest that when the comparison stimulus was presented outside the background, perceptual rescaling mechanisms are not triggered for the

stimulus placed at a location where the pictorial depth cues signal a closer distance (Cretenoud et al., 2020; Yildiz et al., 2019, 2021a). Perhaps because of this, participants in these studies may have spent more time on the comparison stimulus, presented outside the background, rather than the bottom standard stimulus, presented inside the background.

Finally, one might also expect to find increases in the strength of the illusion during development. As demonstrated in the first section, studies examining the role of age-related differences in the strength of the Ponzo-like illusions have provided mixed evidence. For example, Brislin (1974) demonstrated that the magnitude of the field and railroad illusions increased with age in Pennsylvanian but not Guamanian participants when both the comparison and standard stimuli were presented over the same illusory background (the so-called direct comparison task). Contrarily, Cretenoud et al. (2020) demonstrated that there was a slight decline in the magnitude of railroad illusion with age when the comparison stimulus was presented outside the illusory background (the so-called indirect comparison task). Priors about how two physically identical objects subtend different visual angles on the retina could explain why the strength of the Ponzo-variant illusions increases in the direct comparison tasks while it decreases in the indirect comparison tasks with age. As mentioned before, perhaps, the visual system uses the same prior in the direct and indirect comparison tasks until a certain developmental stage. As the brain learns to use a more proper prior for rescaling apparently far stimulus' size, a decline in the strength of the illusion with age is observed when an indirect comparison task is used.

In this Bayesian-motivated reconceptualization of the misapplied size constancy theory, we posit that the classic Ponzo illusion is driven by priors that are possibly innate or acquired at early development stages. There is a possibility that the mechanisms described in the assimilation and tilt-constancy theories of the classic Ponzo illusion help the brain acquire these priors. Therefore, assimilation, tilt constancy, and misapplied size constancy may all function simultaneously. In the later stage, these priors help the brain to selectively process environmental statistics that are essential to forming prior probability density functions. Once the prior probability functions are formed, the relative certainty of the prior information becomes influential on the illusion's strength. Supporting this line of reasoning, many cross-cultural studies suggest that the relative certainty of the prior information increases with the number of pictorial depth cues and affects the strength of the illusion only if the observer has previous experience with pictorial depth cues.

To conclude, the present reconceptualization helps to reconcile inconsistent findings on the Ponzo-variant illusion and explains all Ponzo-like illusions with prior information and prediction errors. The proposed model goes beyond



Gregory's misapplied size constancy theory and underlines differences between the underlying mechanisms of the Ponzo-variant illusions and the classic Ponzo illusion.

**Data availability** Not applicable

## Declarations

This work was supported by a La Trobe University Postgraduate Research Scholarship and a La Trobe University Full Fee Research Scholarship awarded to G.Y.Y.

**Conflict of interest statement:** The authors declare no competing interests.

**Ethics approval** Not applicable

**Consent to participate** Not applicable

**Consent for publication** Not applicable

**Code availability** Not applicable

## References

- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, 7(10), 1057–1058. <https://doi.org/10.1038/nn1312>
- Andres, E., McKyton, A., Ben-Zion, I., & Zohary, E. (2017). Size constancy following long-term visual deprivation. *Current Biology*, 27(14), 696–697. <https://doi.org/10.1016/j.cub.2017.05.071>
- Blakemore, C., Carpenter, R. H., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, 228, 37–39.
- Brislin, R. (1974). The Ponzo illusion: Additional cues, age, orientation, and culture. *Journal of Cross-Cultural Psychology*, 5, 139–161.
- Brislin, R. W., & Leibowitz, H. W. (1970). The effect of separation between test and comparison objects on size constancy at various age-levels. *The American Journal of Psychology*, 83(3), 372–376.
- Brown, J. M., Breitmeyer, B. G., Hale, R. G., & Plummer, R. W. (2018). Contrast sensitivity indicates processing level of visual illusions. *Journal of Experimental Psychology—Human Perception and Performance*, 44(10), 1557–1566. <https://doi.org/10.1037/xhp0000554>
- Chater, N., Clark, A., Goldsmith, J. A., & Perfors, A. (2015). *Empiricism and language learnability*. Oxford University Press.
- Chen, L. H., Qiao, C. Y., Wang, Y., & Jiang, Y. (2018). Subconscious processing reveals dissociable contextual modulations of visual size perception. *Cognition*, 180, 259–267. <https://doi.org/10.1016/j.cognition.2018.07.014>
- Chevrier, J., & Delorme, A. (1983). Depth-perception in pandoras box and size illusion—Evolution with age. *Perception*, 12(2), 177–185. <https://doi.org/10.1068/p120177>
- Chokron, S., & De Agostini, M. (2000). Reading habits influence aesthetic preference. *Cognitive Brain Research*, 10, 45–49.
- Cretenoud, A. F., Karimpur, H., Grzeczowski, L., Francis, G., Hamburger, K., & Herzog, M. H. (2019). Factors underlying visual illusions are illusion-specific but not feature-specific. *Journal of Vision*, 19(14), 1–21. <https://doi.org/10.1167/19.14.12>
- Cretenoud, A. F., Grzeczowski, L., Bertamini, M., & Herzog, M. H. (2020). Individual differences in the Müller-Lyer and Ponzo illusions are stable across different contexts. *Journal of Vision*, 20(6), 4. <https://doi.org/10.1167/jov.20.6.4>
- Cretenoud, A. F., Grzeczowski, L., Kunchulia, M., & Herzog, M. H. (2021). Individual differences in the perception of visual illusions are stable across eyes, time, and measurement methods. *Journal of Vision*, 21(5), 26. <https://doi.org/10.1167/jov.21.5.26>
- Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, 398(6724), 203–204. <https://doi.org/10.1038/18335>
- de Fockert, J., Davidoff, J., Fagot, J., Parron, C., & Goldstein, J. (2007). More accurate size contrast judgments in the Ebbinghaus Illusion by a remote culture. *Journal of Experimental Psychology—Human Perception and Performance*, 33(3), 738–742. <https://doi.org/10.1037/0096-1523.33.3.738>
- Deregowski, J. B. (1968). Difficulties in pictorial depth perception in Africa. *British Journal of Psychology*, 59(3), 195–204. <https://doi.org/10.1111/j.2044-8295.1968.tb01133.x>
- Deregowski, J. B. (1989). Real space and represented space: Cross-cultural perspectives. *Behavioral and Brain Sciences*, 12(1), 51–119. <https://doi.org/10.1017/S0140525X00024286>
- Deregowski, J. B. (2017). Cross-cultural studies of illusions. In A. G. Shapiro & D. Todorović (Eds.), *The Oxford compendium of visual illusions* (pp. 38–53). Oxford University Press.
- Dobias, J. J., Papathomas, T. V., & Sarwate, A. (2016). Ponzo's illusion in 3D: Perspective gradients dominate differences in retinal size. *Multisensory Research*, 29(4/5), 421–438. <https://doi.org/10.1163/22134808-00002522>
- Doherty, M. J., Tsuji, H., & Phillips, W. A. (2008). The context sensitivity of visual size perception varies across cultures. *Perception*, 37(9), 1426–1433. <https://doi.org/10.1068/p5946>
- Fine, I., Wade, A. R., Brewer, A. A., May, M. G., Goodman, D. F., Boynton, G. M., Wandell, B. A., & MacLeod, D. I. A. (2003). Long-term deprivation affects visual perception and cortex. *Nature Neuroscience*, 6(9), 915–916. <https://doi.org/10.1038/nn1102>
- Fineman, M. B. (1981). Complexity of context and orientation of figure in the corridor illusion. *Perceptual and Motor Skills*, 53(1), 11–14. <https://doi.org/10.2466/pms.1981.53.1.11>
- Fineman, M. B., & Carlson, J. (1973). Comparison of Ponzo illusion with a textural analog. *Perception & Psychophysics*, 14(1), 31–33. <https://doi.org/10.3758/Bf03198612>
- Fisher, G. H. (1968a). An experimental and theoretical appraisal of the inappropriate size-depth theories of illusions. *British Journal of Psychology*, 59(4), 373–383. <https://doi.org/10.1111/j.2044-8295.1968.tb01152.x>
- Fisher, G. H. (1968b). Gradients of distortion seen in the context of the Ponzo illusion and other contours. *Quarterly Journal of Experimental Psychology*, 20(2), 212–217. <https://doi.org/10.1080/14640746808400153>
- Fisher, G. H. (1968c). Illusions and size-constancy. *American Journal of Psychology*, 81(1), 2–20.
- Fisher, G. H. (1969). Towards a new explanation for the geometrical illusions: I. The properties of contours which induce illusory distortion. *British Journal of Psychology*, 60(2), 179–185. <https://doi.org/10.1111/j.2044-8295.1969.tb01190.x>
- Fisher, G. H. (1970). An experimental and theoretical appraisal of the perspective and size-constancy theories of illusions. *Quarterly Journal of Experimental Psychology*, 22(4), 631–652. <https://doi.org/10.1080/14640747008401942>
- Fisher, G. H. (1973). Towards a new explanation for the geometrical illusions: II. Apparent depth or contour proximity? *British Journal of Psychology*, 64(4), 607–621. <https://doi.org/10.1111/j.2044-8295.1973.tb01388.x>
- Friedrich, T. E., & Elias, L. J. (2014). Behavioural asymmetries on the greyscales task: The influence of native reading direction. *Cult Brain*, 2(2), 161–172. <https://doi.org/10.1007/s40167-014-0019-3>

- Gandhi, T., Kalia, A., Ganesh, S., & Sinha, P. (2015). Immediate susceptibility to visual illusions after sight onset. *Current Biology*, 25(9), R358–R359. <https://doi.org/10.1016/j.cub.2015.03.005>
- Gardner, R. W., & Long, R. I. (1960). Errors of the standard and illusion effects with inverted-T. *Perceptual and Motor Skills*, 10, 47–54.
- Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines: I. Quantitative studies. *Journal of Experimental Psychology*, 20, 453–467. <https://doi.org/10.1037/h0059826>
- Gilliam, B. (1973). The nature of size scaling in the Ponzo and related illusions. *Perception & Psychophysics* 14, 353–357.
- Girgus, J. S., & Coren, S. (1982). Assimilation and contrast illusions - Differences in plasticity. *Perception & Psychophysics*, 32(6), 555–561. <https://doi.org/10.3758/Bf03204210>
- Granrud, C. E. (2006). Size constancy in infants: 4-month-olds' responses to physical versus retinal image size. *Journal of Experimental Psychology—Human Perception and Performance*, 32(6), 1398–1404. <https://doi.org/10.1037/0096-1523.32.6.1398>
- Green, F. A. (1972). The Relationship between Visual Illusions and Cues to Distance (Unpublished doctoral dissertation). Durham University.
- Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature*, 199, 678–680. <https://doi.org/10.1038/199678a0>
- Gregory, R. L. (1968). Perceptual illusions and brain models. *Proceedings of the Royal Society B: Biological Sciences*, 171(1024), 279–296. <https://doi.org/10.1098/rspb.1968.0071>
- Gregory, R. L. (1998). *Eye and brain: The psychology of seeing* (5th ed.). Oxford University Press.
- Gregory, R. L. (2005). Editorial essay: Bayes Window (1). *Perception*, 34(12), 1421–1422. <https://doi.org/10.1068/p3412ed>
- Gregory, R. L. (2006a). Bayes Window (2): Editorial essay. *Perception*, 35, 143–144. <https://doi.org/10.1068/p3502ed>
- Gregory, R. L. (2006b). Bayes Window (4): Table of illusions. Editorial essay. *Perception*, 35, 431–432. <https://doi.org/10.1068/p3504ed>
- Gregory, R. L. (2009). *Seeing through illusions*. Oxford University Press.
- Grzeczkowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research*, 141, 282–292.
- Hadad, B. S. (2018). Developmental trends in susceptibility to perceptual illusions: Not all illusions are created equal. *Attention, Perception, & Psychophysics*, 80(6), 1619–1628. <https://doi.org/10.3758/s13414-018-1529-4>
- Hemker, L., & Kavsek, M. (2010). The relative contribution of relative height, linear perspective, and texture gradients to pictorial depth perception in 7-month-old infants. *Perception*, 39(11), 1476–1490. <https://doi.org/10.1068/p6682>
- Howe, C. Q., & Purves, D. (2005). *Perceiving geometry: Geometrical illusions explained by natural scene statistics*. Springer.
- Howe, C. Q., Lotto, R. B., & Purves, D. (2006). Comparison of Bayesian and empirical ranking approaches to visual perception. *Journal of Theoretical Biology*, 241(4), 866–875. <https://doi.org/10.1016/j.jtbi.2006.01.017>
- Hudson, W. (1960). Pictorial depth-perception in sub-cultural groups in Africa. *Journal of Social Psychology*, 52(2), 183–208. <https://doi.org/10.1080/00224545.1960.9922077>
- Jaeger, T., Treiber, F., & Pollack, R. H. (1980). Effect of lightness contrast on Ponzo illusions. *Bulletin of the Psychonomic Society*, 15(1), 1–4.
- Jahoda, G., & McGurk, H. (1974). Development of pictorial depth perception: Cross-cultural replications. *Child Dev*, 45(4), 1042–1047.
- Jenkin, N., & Feallock, S. M. (1960). Developmental and intellectual processes in size-distance judgment. *American Journal of Psychology*, 73(2), 268–273. <https://doi.org/10.2307/1419904>
- Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The flash-lag effect as a motion-based predictive shift. *PLOS Computational Biology*, 13(1). <https://doi.org/10.1371/journal.pcbi.1005068>
- Kilbride, P. L., & Leibowitz, H. W. (1975). Factors affecting the magnitude of the Ponzo perspective illusion among the Baganda. *Perception & Psychophysics*, 17(6), 543–548.
- Kilbride, P. L., & Robbins, M. C. (1968). Linear perspective pictorial depth perception and education among Baganda. *Perceptual and Motor Skills*, 27(2). <https://doi.org/10.2466/pms.1968.27.2.601>
- Knill, D. C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Research*, 43(7), 831–854. [https://doi.org/10.1016/S0042-6989\(03\)00003-8](https://doi.org/10.1016/S0042-6989(03)00003-8)
- Lazar, I. (1964). Recovery from early blindness - A case-study - Gregory, R. L., Wallace, J. *Contemporary Psychology*, 9(7).
- Leibowitz, H. W., & Judisch, J. M. (1967). The relation between age and the magnitude of the Ponzo illusion. *American Journal of Psychology*, 80(1), 105–109.
- Leibowitz, H. W., & Pick, H. A. (1972). Cross-cultural and educational aspects of Ponzo perspective illusion. *Perception & Psychophysics*, 12(5). <https://doi.org/10.3758/Bf03205856>
- Leibowitz, H. W., Pollard, S. W., & Dickson, D. (1967). Monocular and binocular size-matching as a function of distance at various age-levels. *American Journal of Psychology*, 80(2). <https://doi.org/10.2307/1420986>
- Leibowitz, H., Brislin, R., Perlmutter, L., & Hennessy, R. (1969). Ponzo perspective illusion as a manifestation of space perception. *Science*, 166(3909), 1174–1176. <https://doi.org/10.1126/science.166.3909.1174>
- Maurer, D., Lewis, T. L., Brent, H. P., & Levin, A. V. (1999). Rapid improvement in the acuity of infants after visual input. *Science*, 286(5437), 108–110. <https://doi.org/10.1126/science.286.5437.108>
- McGurk, H., & Jahoda, G. (1975). Pictorial depth perception by children in Scotland and Ghana. *Journal of Cross-Cultural Psychology*, 6(3), 279–296. <https://doi.org/10.1177/002202217563002>
- Miller, R. J. (1973). Cross-cultural research in the perception of pictorial materials. *Psychological Bulletin*, 80(2), 135–150. <https://doi.org/10.1037/h0034739>
- Müller-Lyer, F. C. (1889). Optische urteilstäuschungen [Optical misjudgments]. *Archiv für Physiologie Suppl.*, 263–270.
- Mundy-Castle, A. C. (1966). Pictorial depth perception in Ghanaian children. *International Journal of Psychology*, 1(4), 289–300.
- Newman, C. V. (1969). Children's size judgments in a picture with suggested depth. *Nature*, 223(5204). <https://doi.org/10.1038/223418a0>
- Piaget, M. (1969). *The mechanisms of perception*. Routledge.
- Plude, D. J., Milberg, W. P., & Cerella, J. (1986). Age-differences in depicting and perceiving tridimensionality in simple line drawings. *Experimental Aging Research*, 12(4), 221–225.
- Ponzo, M. (1910). *Intorno ad alcune illusioni nel campo delle sensazioni tattili, sull'illusione di Aristotele e fenomeni analoghi*. Leipzig, Germany: Wilhelm Engelmann.
- Predebon, J. (1985). Age trends in the Muller-Lyer and Ponzo illusions. *British Journal of Developmental Psychology*, 3(Mar), 99–103. <https://doi.org/10.1111/j.2044-835X.1985.tb00960.x>
- Pressey, A. W. (1972). The assimilation theory of geometric illusions: An additional postulate. *Perception & Psychophysics*, 11(1A), 28–30.
- Pressey, A. W. (1974a). Age-changes in Ponzo and filled-space illusions. *Perception & Psychophysics*, 15(2), 315–319. <https://doi.org/10.3758/Bf03213950>
- Pressey, A. W. (1974b). Evidence for the role of attentive fields in the perception of illusions. *Quarterly Journal of Experimental Psychology*, 26(Pt. 3), 464–471. <https://doi.org/10.1080/14640747408400435>
- Pressey, A. W. (1987). Psychophysical methods and life-span changes in visual illusions. *Perceptual and Motor Skills*, 65(1), 83–87. <https://doi.org/10.2466/pms.1987.65.1.83>

- Pressey, A. W. (2013). Clashing interpretations of Ponzo's illusion: Reconsideration of assimilation versus tilt-constancy theories. *Comprehensive Psychology*, 2(14). <https://doi.org/10.2466/24.27.CP.2.14>
- Pressey, A. W., & Epp, D. (1992). Spatial attention in Ponzo-like patterns. *Perception & Psychophysics*, 52(2), 211–221. <https://doi.org/10.3758/bf03206774>
- Pressey, A. W., Butchard, N., & Scrivner, L. (1971). Assimilation theory and the Ponzo illusion: Quantitative predictions. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 25(6), 486–497. <https://doi.org/10.1037/h0082409>
- Prinzmetal, W., & Beck, D. M. (2001). The tilt-consistency theory of visual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 206–217. <https://doi.org/10.1037/0096-1523.27.1.206>
- Prinzmetal, W., Shimamura, A. P., & Mikolinski, M. (2001). The Ponzo illusion and the perception of orientation. *Perception & Psychophysics*, 63(1), 99–114. <https://doi.org/10.3758/bf03200506>
- Qian, J., Liu, S., & Lei, Q. (2016). Illusory distance modulates perceived size of afterimage despite the disappearance of depth cues. *PLOS ONE*, 11(7). <https://doi.org/10.1371/journal.pone.0159228>
- Quina, K., & Pollack, R. H. (1972). Effects of test line position and age on the magnitude of the Ponzo illusion. *Perception & Psychophysics*, 12(2), 435–439. <https://doi.org/10.3758/BF03203036>
- Rima, S., Khalil, C., Cottureau, B. R., Trotter, Y., & Durand, J. B. (2019). Asymmetry of pictorial space: A cultural phenomenon. *Journal of Vision*, 19(4). <https://doi.org/10.1167/19.4.22>
- Roncato, S., Parlangeli, O., & Farfaneti, G. (1998). Effects of contextual and local factors on Ponzo illusion magnitude. *Psychological Research—Psychologische Forschung*, 61(2), 119–124. <https://doi.org/10.1007/s004260050018>
- Sander, F. (1926). Optische täuschungen und physiologie. *Neue Psychologische Studien 1*, 159–166.
- Schiffman, H. R., & Thompson, J. G. (1978). The role of apparent depth and context in the perception of the Ponzo illusion. *Perception*, 7(1), 47–50. <https://doi.org/10.1068/p070047>
- Schmidt, F., & Haberkamp, A. (2016). Temporal processing characteristics of the Ponzo illusion. *Psychological Research*, 80(2), 273–285. <https://doi.org/10.1007/s00426-015-0659-8>
- Segall, M. H., Campbell, D. T., & Herskovits, M. J. (1966). The influence of culture on visual perception. In H. Toch & C. Smith (Eds.), *Social Perception*. Indianapolis: .
- Seydell, A., Knill, D. C., & Trommershäuser, J. (2011). Priors and learning in cue integration. In J. Trommershäuser, K. Kording, & M. S. Landy (Eds.), *Sensory cue integration*. Oxford University Press.
- Shen, M. W., Xu, H. K., Zhang, H. H., Shui, R. D., Zhang, M., & Zhou, J. F. (2015). The working memory Ponzo illusion: Involuntary integration of visuospatial information stored in visual working memory. *Cognition*, 141, 26–35. <https://doi.org/10.1016/j.cognition.2015.04.010>
- Simpson, T., Carruthers, P., Laurence, S., & Stich, S. (2005). Nativism past and present. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind* (pp. 3–19). : Oxford University Press.
- Slater, A., Mattock, A., & Brown, E. (1990). Size constancy at birth—Newborn-infants' responses to retinal and real size. *Journal of Experimental Child Psychology*, 49(2), 314–322. [https://doi.org/10.1016/0022-0965\(90\)90061-C](https://doi.org/10.1016/0022-0965(90)90061-C)
- Song, C., Schwarzkopf, D. S., & Rees, G. (2011). Interocular induction of illusory size perception. *BMC Neurosci*, 12, 27. <https://doi.org/10.1186/1471-2202-12-27>
- Sperandio, I., & Chouinard, P. A. (2015). The mechanisms of size constancy. *Multisensory Research*, 28(3/4), 253–283.
- Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nat Neurosci*, 15(4), 540–542. <https://doi.org/10.1038/nn.3069>
- Teller, D. Y. (1997). First glances: The vision of infants—The Friedenwald Lecture. *Investigative Ophthalmology & Visual Science*, 38(11), 2183–2203.
- Wagner, D. A. (1977). Ontogeny of the Ponzo illusion: Effects of age, schooling, and environment. *International Journal of Psychology*, 12(3).
- Weiss, Y., Simoncelli, E. P., Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598–604. <https://doi.org/10.1038/nm0602-858>
- Wilcox, B. L., & Teghtsoonian, M. (1971). The control of relative size by pictorial depth cues in children and adults. *Journal of Experimental Child Psychology*, 11(3), 413–429. [https://doi.org/10.1016/0022-0965\(71\)90046-4](https://doi.org/10.1016/0022-0965(71)90046-4)
- Wohlwill, J. F. (1962). Perspective illusion perceived size and distance in fields varying in suggested depth, in children and adults. *Journal of Experimental Psychology*, 64(3). <https://doi.org/10.1037/h0041720>
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2019). The contribution of linear perspective cues and texture gradients in the perceptual rescaling of stimuli inside a Ponzo illusion corridor. *PLOS ONE*, 14(10). <https://doi.org/10.1371/journal.pone.0223583>
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2021a). The conceptual understanding of depth rather than the low-level processing of spatial frequencies drives the corridor illusion. *Vision Research*, 181, 21–31. <https://doi.org/10.1016/j.visres.2020.11.012>
- Yildiz, G. Y., Sperandio, I., Kettle, C., & Chouinard, P. A. (2021b). Interocular transfer effects of linear perspective cues and texture gradients in the perceptual rescaling of size. *Vision Research*, 179, 19–33. <https://doi.org/10.1016/j.visres.2020.11.005>
- Yonas, A., Cleaves, W. T., & Pettersen, L. (1978). Development of sensitivity to pictorial depth. *Science*, 200(4337), 77–79. <https://doi.org/10.1126/science.635576>
- Yonas, A., Elieff, C. A., & Arterberry, M. E. (2002). Emergence of sensitivity to pictorial depth cues: Charting development in individual infants. *Infant Behavior & Development*, 25(4), 495–514. [https://doi.org/10.1016/S0163-6383\(02\)00147-9](https://doi.org/10.1016/S0163-6383(02)00147-9)
- Zeigler, H. P., & Leibowitz, H. (1957). Apparent visual size as a function of distance for children and adults. *American Journal of Psychology*, 70(1), 106–109.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.