



## Transitivity of visual sameness

Błażej Skrzypulec<sup>1</sup> 

Received: 23 January 2017 / Accepted: 20 June 2018  
© The Author(s) 2018

### Abstract

The way in which vision represents objects as being the same despite movement and qualitative changes has been extensively investigated in contemporary psychology. However, the formal properties of the visual sameness relation are still unclear, for example, whether it is an identity-like, equivalence relation. The paper concerns one aspect of this problem: the transitivity of visual sameness. Results obtained by using different experimental paradigms are analysed, in particular studies using streaming/bouncing stimuli, multiple object tracking experiments and investigations concerning object-specific preview benefit, and it is argued that the transitive interpretation of visual sameness is the most plausible given the current stage of knowledge. What is more, it is claimed that the way in which visual sameness is represented suggests that in some cases it should be characterized as a “primitive sameness”, similarly as in philosophical theories postulating “thisness”.

**Keywords** Visual sameness · Transitivity · Identity · Vision · Perception · Multiple object tracking

In contemporary cognitive psychology, it is widely believed that human vision represents the numerical sameness of objects, and that objects can be re-identified as being the same after movement and changes in their features (Scholl 2007; Pylyshyn 2007). Researchers have obtained important results regarding conditions for representing sameness, concerning, in particular, the significant role of spatiotemporal continuity and cohesion (Kahneman et al. 1992). Furthermore, the mechanisms responsible for object re-identification have been studied in order to recognise the role of attentive and preattentive processes (Pylyshyn 2001; Oksama and Hyönä 2004).

Such psychological results suggest that human vision is able not only to represent that entities are qualitatively the same (e.g., like two indistinguishable shades of a colour), but also that an entity  $x$  is numerically the same individual as an entity  $y$ . In

---

✉ Błażej Skrzypulec  
blazej.skrzypulec@gmail.com

<sup>1</sup> Institute of Philosophy and Sociology, Polish Academy of Sciences, Warsaw, Poland

addition, the occurrence of numeric visual sameness does not depend on the occurrence of visual qualitative sameness, since objects with differing properties can be visually represented as being numerically the same (e.g., continuously moving objects that change their colour). This is an important feature of a “visual ontology”, or in other words, the ontology of the environment, as it is represented by vision. While it may be asked whether the visual ontology of sameness matches the actual ontology of physical things, in this paper I am not concerned with this issue, I investigate only how numeric sameness is visually represented.

Despite the extensive psychological works regarding mechanisms responsible for representing sameness, the relation of visual numeric sameness itself has not attracted as much attention. As we know from philosophical works, the relation of numeric sameness can be characterised in various ways, by ascribing different properties to it (Chisholm 1969). One of the most important questions is whether this relation is one of classical identity, and thus whether it is an equivalence relation. The positive answer for this question would mean that visually representing entities as numerically the same, can be equated with representing them as numerically identical. In this paper I consider one aspect of this question: the transitivity of numeric visual sameness.

Formal properties of relations have already been investigated by psychologists. In his classic research, Tversky (1977) tried to establish whether people treat similarity as a symmetric relation or not. There are also works in which the transitivity of numeric sameness is considered, however they do not investigate visual representations, but people’s judgments based on stories they have read (Rips 2011). Furthermore, philosophers have been discussing the transitivity of visual sameness, but have been focusing on the qualitative, and not the numeric sameness. In particular, a perceptual version of the sorites paradox has been investigated by discussing whether, in a series of perceptual qualities such that each element is indistinguishable from the subsequent element, distant elements are also indistinguishable (see Hardin 1988; Raffman 2000; Graff 2001; Mills 2002; Chuard 2007; Williamson 2013). Here, I analyse empirical results concerning visual representations of numeric sameness and explore what we can infer from them about the transitivity of the considered relation. This can be regarded as a first step towards constructing a theory of numeric visual sameness.

The paper starts by introducing the notion of a “splitting-like situation”. Such situations are suitable for testing the transitivity of numeric sameness (Sect. 1). Subsequently, I analyse the results of investigations using streaming/bouncing stimuli (Sect. 2), measuring Object-Specific Preview Benefit (Sect. 3), and applying the Multiple Object Tracking paradigm (Sect. 4) in order to check whether they determine the transitivity of numeric visual sameness.<sup>1</sup> Finally (Sects. 5 and 6), I describe the results of conducted analyses. I argue that according to the current state of research, the transitive interpretation of numeric visual sameness is more plausible. Furthermore, I claim that the way in which numeric visual sameness is represented suggests that

---

<sup>1</sup> There are also interesting studies on infants and primates in which splitting-like situations are used (e.g., Xu and Carey 1996; Hood and Bloom 2008; Cacchione and Call 2011). However, I do not discuss them here as what they investigate is not how vision represents sameness, but how sameness is represented as a result of a cooperation between perceptual and higher-level, cognitive processes. Nevertheless, the results presented in this paper can be interpreted as concerning a perceptual aspect of sameness representations, which also influences the way that sameness is represented by cognitive mechanisms.

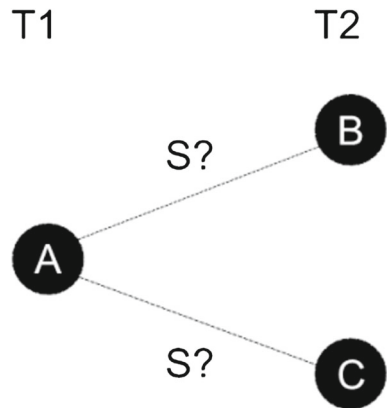
in some cases it should be characterised as a “primitive sameness” (Sect. 5). Further in the paper, if not mentioned otherwise, by “visual sameness” I will mean “numeric visual sameness”.

## 1 How to test transitivity?

Before starting the investigations concerning the transitivity of visual sameness, it is worth specifying the exact problem I have in mind, as there are at least several philosophically interesting issues connected with perceptual representations of sameness. First, I do not investigate the features that usually determine whether an object at one moment is visually represented as being the same as an object perceived at a subsequent moment. In this case, I follow the psychological consensus, that changes concerning topological properties and spatiotemporal continuity are far more likely to break the represented sameness, than changes in properties like colour, size, or position (e.g., von Marle and Scholl 2003; Mitroff and Alvarez 2007; Hollingworth and Franconeri 2009; Moore et al. 2010). Second, I do not make any assumptions regarding the relation between representing visual sameness and phenomenal character of visual experiences. There may be a phenomenal change occurring when sameness is represented. For instance, participants in some experiments report “feeling” that an object is the same (Natsoulas 1992), but it is difficult to establish whether such an effect is related to representing sameness and not, for example, related to representing a specific pattern of spatiotemporal relations. Third, my investigations concern the numeric and not the qualitative sameness. In other words, they are about representing whether an object  $x$  is the same individual as an object  $y$ , and not whether  $x$  and  $y$  are indistinguishable in virtue of having the same properties like colour, shape, etc. In the case of human vision, as suggested by psychological results (see Pylyshyn 2007; Scholl 2007 for a review), representing numeric sameness is independent from representing qualitative sameness. Objects represented as qualitatively the same may be represented as numerically distinct if, for instance, there is a temporal discontinuity between presentation of the first and the second object. Similarly, an entity perceptually recognised as a single object may be represented as having different qualitative properties at various moments of its existence, as not all qualitative changes break a visual representation of the numeric sameness.

More specifically, I analyse the transitivity of visual sameness in the context of a continuous observation, performed by a single subject under normal processing conditions. This means that the cases under discussion involve a single subject, who visually represents an object  $x$  as being (or not being) numerically the same individual as an object  $y$ . By “normal processing condition” I mean a requirement that there are enough cognitive resources to establish visual sameness relations. For instance, such a requirement may not be fulfilled if a subject is attending to some other task, and there are not enough resources left to represent sameness relations between  $x$  and  $y$ . This requirement also excludes cases in which the temporal distance between  $x$  and  $y$  is too long, such that visual mechanisms are no longer able to make relevant comparisons between them. Furthermore, the observation of objects must be continuous, as in cases of significant discontinuity, like judging whether an object seen an hour ago

Fig. 1 A splitting-like situation



is numerically the same as an object seen now; the sameness is established mainly as a result of postperceptual processes. Relying on the above constraints, the transitivity of numeric visual sameness, as understood in this paper, can be characterised as follows: the numeric visual sameness relation, represented by a single subject during a continuous observation under normal processing conditions, is transitive if and only if, if an object A is visually represented as being numerically the same as an object B, and B is visually represented as being numerically the same as an object C, then A is visually represented as being numerically the same as C.

In philosophical works the transitivity of sameness is tested by constructing, usually in the form of a thought-experiment, a splitting-like situation (see Fig. 1 and Parfit 1971 for a classical example). In such a situation there is an object A existing at some moment  $T_1$  and two objects, B and C, existing at a subsequent moment  $T_2$ . Further, both pairs A/B and A/C satisfy conditions  $\{c_1, c_2, \dots\}$  that in ordinary circumstances are necessary and sufficient for the diachronic sameness between objects of the considered kind.<sup>2</sup> In addition, objects B and C are numerically distinct.

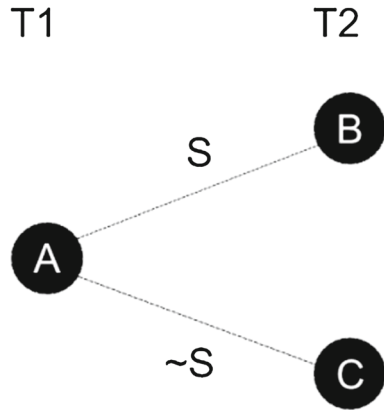
In other words, in a splitting-like situation there are two equally good candidates for being the same as the object A. Given this, the splitting-like situation can be resolved in three general ways:

### 1.1 Unique sameness

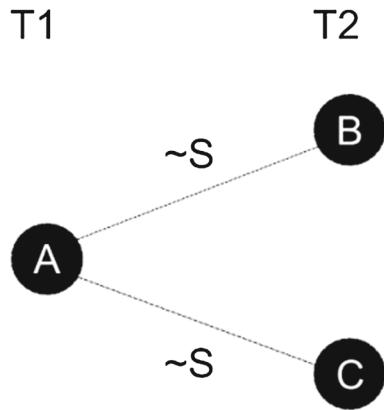
The object A is the same as exactly one object at  $T_2$  (Fig. 2). This result is consistent with the transitivity of sameness. However, if sameness occurs only between A and one object at  $T_2$ , then the set  $\{c_1, c_2, \dots\}$  of sameness conditions should be modified, because there is probably an aspect in which the objects B and C differ determining

<sup>2</sup> I assume an endurantist view on visual persistence and diachronic sameness, according to which, objects can be “wholly present” at different moments, and a sameness relation may occur between such objects. From the perspective of an alternative, perdurantist view, the visual sameness would occur between temporal parts of a single, temporally extended object (see Lewis 1986, p. 202 for a classical source of this distinction). This assumption is motivated by the wide-held conviction among philosophers of a perception that human vision represents persistence in the endurantist way (possibly in contrast to human audition, see Nudds 2014; O’Callaghan 2014).

**Fig. 2** Unique sameness. Object A is the same with exactly one object at  $T_2$



**Fig. 3** End of sameness. Object A is not the same as any of the objects at  $T_2$

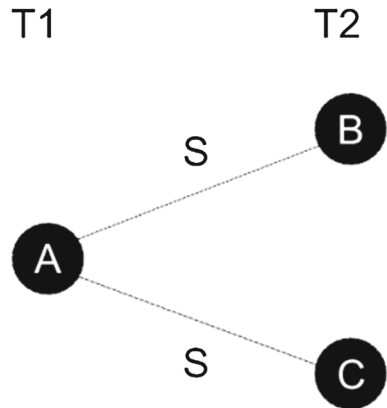


which of them is the same as A. Without including this additional aspect, the joint fulfilment of the conditions  $\{c_1, c_2, \dots\}$  is not sufficient for the sameness of objects.

### 1.2 End of sameness

The object A is not the same as any of the objects at  $T_2$  (Fig. 3). In this case a splitting-like situation breaks the sameness between objects. This result is also consistent with the transitivity of sameness. In addition, there is no need to modify the set  $\{c_1, c_2, \dots\}$  of sameness conditions in any complicated way. To formulate a diachronic sameness criterion it is enough to state that an object  $x$  is the same as an object  $y$  if and only if (1) they satisfy conditions  $\{c_1, c_2, \dots\}$  and (2) there is no splitting-like situation involving  $x$  and  $y$ .

**Fig. 4** Double sameness. Object A is the same as both objects at  $T_2$



### 1.3 Double sameness

The object A is the same as both objects at  $T_2$  (see Fig. 4). In this case the sameness cannot be characterised as transitive. If sameness were transitive,<sup>3</sup> then objects B and C would also be the same. However, this leads to a contradiction, since in a splitting-like situation they are different objects. This result does not lead to modification of the set  $\{c_1, c_2, \dots\}$ , as satisfying these conditions is sufficient for the sameness of objects.

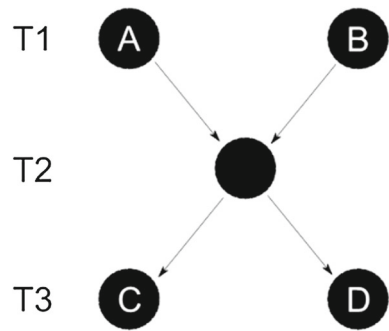
Splitting-like situations constitute a test for the transitivity of sameness, since the occurrence of “Double sameness” falsifies the hypothesis that sameness is transitive.

It should be noted that the exact number of objects involved, i.e. one object at  $T_1$  and two objects at  $T_2$ , is not essential for a splitting-like situation. What is important is the presence of an ambiguity concerning sameness between at least one object at an earlier moment and at least two objects at a subsequent moment. Furthermore, a splitting-like situation does not demand that objects overlap (see Fig. 8 on p. 17 for an example), it is enough that objects are close to each other at subsequent moments. Re-identifying and tracking moving objects through time, and predicting their future spatial features, is a difficult task, which often involves mistakes (Intriligator and Cavanagh 2001; Pylyshyn 2004; Keane and Pylyshyn 2006). Hence, in situations when several objects are close to each other, it may be ambiguous how to establish sameness relations between entities at earlier and later moments.

In addition, one splitting-like situation may be weaker or stronger relative to another one. A splitting-like situation X is stronger than a splitting-situation Y if and only if a set of sameness conditions connected with Y is a subset of a set of sameness conditions connected with X. The notion of stronger and weaker splitting-like situations is important, because stronger splitting-like situations provide a more reliable test for transitivity. If a splitting-like situation is resolved as “Unique sameness” or “End of sameness”, it may still be the case that a stronger splitting-like situation, that is resolved in a way described as “Double sameness” exists. Such a result would falsify

<sup>3</sup> In this paper I do not discuss symmetry of visual sameness: I assume, in accordance with an intuitive conviction, that it is symmetric.

**Fig. 5** Streaming/bouncing stimuli



the hypothesis that sameness is transitive, even if resolutions of weaker splitting-like situations were consistent with transitivity.

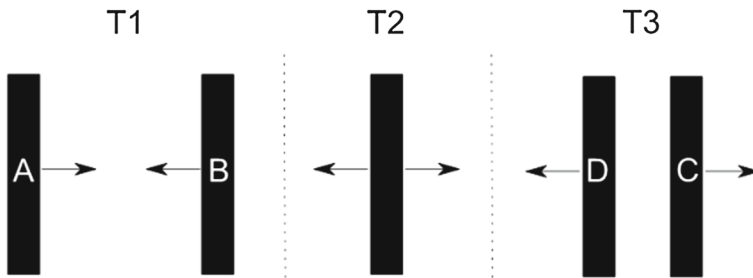
Further, I analyse whether results of investigations into the way in which vision represents sameness, provide information about the patterns of visual sameness occurring in splitting-like situations.

## 2 Streaming/bouncing stimuli

Probably the best-known type of research concerning visual splitting-like situations uses ambiguous streaming/bouncing stimuli (Julesz 1995, p. 50). In such investigations, participants see two objects that are closing in on each other. At one point their trajectories intersect and the objects completely overlap. After that they continue their movement, becoming spatially separated again (Fig. 5).

The pattern illustrated in Fig. 5 is a form of a splitting-like situation. Each of the objects before the overlap, A and B, is a plausible candidate for being the same as each of the objects after the overlap (C and D). This is because (a) all four objects have the same properties, such as shape and colour, and (b) object A, as well as object B, is spatiotemporally continuous with both objects C and D. If object A is represented as being the same as object D, and object B is represented as being the same as object C, then the stimulus will be perceived as “streaming”, i.e. objects are passing through each other during the moment of overlap. However, if object A is represented as being the same as C and object B is represented as being the same as D, the pattern will be seen as “bouncing”, i.e. objects collide during intersection and change their trajectories.

When the standard version of the streaming/bouncing stimulus is presented, as depicted in Fig. 5, the majority of participants report seeing “streaming” (Mitroff et al. 2005, p. 71). This result may be changed by modifying the stimulus. For example, there are more “bouncing” responses when objects slow down before the overlap (Sekuler and Sekuler 1999). However, it is most important for investigations concerning transitivity that both “streaming” and “bouncing” responses state that each of the objects A and B is the same as exactly one object presented after the overlap. This suggests that such splitting-like situations are resolved in accordance with “Unique sameness”, so their resolution is consistent with the transitivity of visual sameness. Nothing in the participants’ responses suggests that these splitting-like situations break



**Fig. 6** A variant of a streaming/bouncing stimulus

sameness (“End of sameness”) or that A and B are the same as both objects C and D (“Double sameness”).

Nevertheless, the usual version of streaming/bouncing stimuli constitutes a rather weak splitting-like situation. In particular, it is easy to explain why object A is usually represented as being the same as D and not the same as C (and analogously for object B). The visual system is able to predict, to some degree, the future location of attended objects relying on direction of movement and velocity (Atsma et al. 2012; Howe and Holcombe 2012). However, such a prediction cannot account for rapid changes in motion direction, and vision will probably predict that an object will continue to move in a straight line. Hence, in a standard streaming/bouncing pattern, object D will be recognised as being the same as object A, since this identification is coherent with predictions concerning the future position of A. One may suppose that in a stronger splitting-like situation, connected with a streaming/bouncing stimulus, the pattern of sameness would differ—possibly not supporting the transitivity of visual sameness.

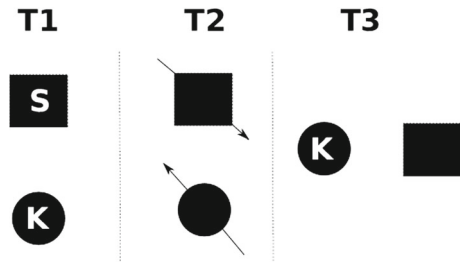
In fact, a description of this stronger splitting-like situation can be found in the literature (Caplovitz et al. 2011). In this case the streaming/bouncing stimulus takes the form of two vertical bars that move horizontally to the centre of the display, overlap each other, and then are disconnected again by moving in the opposite direction (Fig. 6).

The investigators changed some features of the bars to check how they would influence participants’ responses. When both bars had the same features, the dominant response was “streaming,” as in the standard streaming/bouncing stimulus, so object A was represented as being the same as C, and object B as being the same as D (Fig. 6). However, the pattern of responses differed when bars A and D were filled-in by sine-wave textures of differing orientations to B and C. In particular, if the difference in orientation was about 30 degrees, “bouncing” responses were as frequent as “streaming” ones (Caplovitz et al. 2011, p. 8). This is a strong splitting-like situation, as the choice between (a) identifying A with C, and B with D, and (b) identifying A with D, and B with C, seems to be made at random, such that it is difficult to determine the additional conditions that underlie one of these two outcomes.

However, even in the above strong splitting-like situation, each one of the objects before the overlap is represented as being the same as exactly one of the objects after the overlap, no matter whether the response is “streaming” or “bouncing”. Hence, studies concerning streaming/bouncing stimuli support the “Unique sameness” option.



**Fig. 7** An example of a display from an OSPB study



Nevertheless, there are at least two reasons why we should not conclude that the question of the transitivity of visual sameness is settled by the considerations about streaming/bouncing stimuli. First, streaming/bouncing patterns are highly specific, and it cannot be automatically inferred that other visual splitting-like situations are resolved in the same way. Second, one may doubt whether the responses evoked by streaming/bouncing stimuli are mainly determined by the functioning of the visual system, since participants' judgments are likely to be preceded by some reasoning. Because of this, further analyse other splitting-like situations that are not connected with streaming/bouncing stimuli.

The conclusions arrived at by analysing streaming/bouncing cases are analogous to those suggested by the classical studies concerning ambiguous apparent motion. If two objects are presented one after another in proximal locations, people experience a single moving object. In the case of an ambiguous apparent motion display, after a presentation of the first object, two objects are shown, both at an equal distance from the initial item. When such a stimulus is presented, people tend to randomly decide which of the later objects is the same as the earlier one, or else claim that the initial object has split into two entities, so each of the later objects is a continuation of a distinct earlier item (Ullman 1979, pp. 35–40; Dawson 1990). These results suggest the “Unique sameness” solution. It even seems that the visual system tends to duplicate the initial object in order to achieve one-to-one correspondence between earlier and later items. However, as in the case of streaming/bouncing stimuli, it is not clear whether the participants' judgments result only from perceptual processes, and how to exactly interpret phenomenal claims concerning duplication of the initial object (for instance, perhaps it is not a duplication, but an intransitive identification of a single object with two continuants).

### 3 Object-specific preview benefit

An important line of psychological research regarding visual sameness is based on a phenomenon known as Object-Specific Preview Benefit (OSPB). It is observed that a feature is recognised faster if it reappears on the same object on which it was presented earlier even if an object has moved (Kahneman et al. 1992). Displays that are used in the OSPB-related investigations usually consist of three stages (see Fig. 7).

At the first stage two objects are presented with letters ‘K’ and ‘S’ inside. After a while, the letters disappear and both objects move. Finally, when the objects are at

some distance from their original locations, a letter ‘K’ appears again in one of them. If the letter reappears on the same object (see Fig. 7), then it is recognised more rapidly than in the case of its reappearance on the second object.

The OSPB paradigm was directly applied to a splitting-like situation in a study by Mitroff et al. (2004). The display they used contained an object that gradually split into two objects, in a bacteria-like fashion. More specifically, at first there were two circular objects, each with a separate letter inside it (let’s say ‘M’ and ‘T’). After this, the letters disappeared and one of the objects (let’s say the one containing the ‘M’) underwent a gradual division, while the second simply moved straight ahead. Finally, when the two spatially separated objects, arising from the object that underwent the division, were clearly visible, the letter ‘M’, or the letter ‘T’, or some new letter, appeared inside one of them. The task of the participants was to recognise whether the letter that appeared was a new one, or one of those that was presented earlier (‘M’ or ‘T’).

Such a display involves a strong splitting-like situation, since both objects that resulted from division are equally good candidates to be the same as the object before division. They are both spatiotemporally continuous with the earlier object, and share with it properties like shape, colour, and size.

It was observed that if the same letter appeared at the beginning of a trial on the dividing object, and also on one of the resulting objects, it was recognised more quickly than in a case when the letters differed. For instance if, at the beginning, one object bore the letter ‘M’ and another the letter ‘T’, and the object bearing the ‘M’ divided, then ‘M’ and ‘T’, when presented on one of the objects produced by the division, differed in terms of the time needed to recognise that they had already been presented earlier: ‘M’ was recognised more rapidly than ‘T’. Furthermore, the reaction time for ‘M’ was similar for each of the two objects that resulted from the division. This suggests that both objects resulting from division were represented as being the same as the initial object. Indeed, the authors agree that this is the most plausible interpretation of the obtained results (Mitroff et al. 2004, p. 424). Of course, such result supports “Double sameness” and so seem to provide a counter-example against the transitivity of visual sameness.

On the other hand, “End of sameness” seems improbable. If the considered splitting-like situation is resolved along this pattern, then the reaction time should not be significantly different for ‘M’ and ‘T’, since none of the new objects would be represented as being the same as the initial one (which contained ‘M’).

The situation is more complicated with respect to “Unique sameness”. If this resolution were correct, then in some trials one of the objects resulting from division would be treated as being the same as the initial one, while in other trials this would be true only in the case of the second object. Therefore the presence of the letter ‘M’ would be privileged on one object, but on the other one, its recognition would be no more efficient than the recognition of ‘T’. The authors of the study claim that “nothing in the obtained results suggests such a pattern, yet they note that its occurrence would be difficult to detect, so they chose to “remain agnostic about which interpretation is correct” (Mitroff et al. 2004, p. 424), i.e., in terms used in this paper, “Unique sameness” or “Double sameness”. If, at each trial, only one object allowed for a faster recognition of ‘M’, it is probable that the overall OSPB in the considered splitting-

like situation should be weaker, than in displays that do not involve division. In fact, this is what has been observed (Mitroff et al. 2004, p. 423). However, an alternative interpretation of this observation is also available, which does not support “Unique sameness”. Weaker OSPBs may arise from the fact that division cases are more cognitively demanding (Mitroff et al. 2004, p. 423). Overall, while “Double sameness” seems to be the most plausible, the results of the study do not allow to definitely reject the “Unique sameness” interpretation.

Results obtained from studies using streaming/bouncing stimuli, and results from the OSPB division study seem to be in conflict. The first supports “Unique sameness”, which is consistent with the transitivity of visual sameness, while the second is likely to provide a counter-example to transitivity. Nevertheless, they both demonstrate that splitting-like situations do not break visual sameness, and that the “End of sameness” solution is incorrect.

Of course, it should be enough to present one example of the intransitive resolution of a splitting-like situation in order to falsify a general hypothesis that visual sameness is transitive. In addition, in contrast to the case of streaming/bouncing stimuli, in OSPB studies the verdict concerning patterns of sameness does not rely on participants’ judgments, so it is more likely to reveal the way in which vision works, without significant top-down modifications.

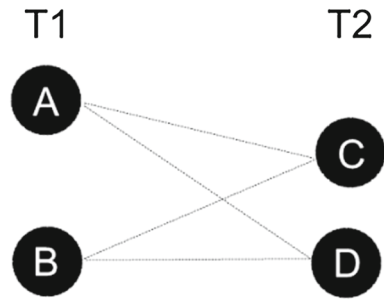
However, the result of this division study is still not conclusive in respect of the question of transitivity. First, as already mentioned, the results do not allow us to decisively reject the “Unique sameness” solution. Second, there is a problem connected with the interpretation of OSPB. It is not clear whether OSPB really depends on the occurrence of visual sameness. Alternatively, it could be that OSPB appears due to spatial relations; for example, in the case of division, it may be connected with spatial continuity between the initial and the two later objects. Another possibility is that the relations connecting objects in the division study are not sameness relations, but some kind of descendance relations. In fact, some authors suggest that OSPB may be a product of low-level operations that recognise spatial proximity (Odic et al. 2012). In this case, the results of the division study do not tell us much about visual sameness. For example, it is possible that the initial object is not represented as being the same as any of the later objects, but OSBP arises in virtue of mere continuity.

In the next section, I consider splitting-like situations from the point of view of the Multiple Object Tracking paradigm. I argue that it is an important tool for testing transitivity.

## 4 Multiple object tracking

In a standard multiple object tracking (MOT) task, introduced by Pylyshyn and Storm (1988), participants are presented with a set of simple objects of the same shape, size, and colour. Initially, some of the presented objects blink a few times—they are designated as targets—while the remaining objects play the role of distractors. Subsequently, all objects start to move in a random fashion, and the participants’ task is to track the targets. After some time the objects stop, and one of them blinks again. Participants are then asked to decide whether this object is a target, or a distractor. The

**Fig. 8** A double splitting-like situation occurring in MOT studies



success rate is usually very high if no more than four targets, among four distractors, are tracked (Pylyshyn 2007, p. 35).

Various MOT-type experiments have been used to investigate factors influencing the visual representation of sameness. If some changes break visual sameness, then the success rate in an MOT task should be lower, because if a target changes in a sameness-breaking way, then after the change it is no longer represented as the same tracked object, but is treated as one of the distractors. Hence, it is not properly recognised at the end of a trial. Analogously, if a change does not worsen the participants' performance, then it may be plausibly stated that an object before and after that change is still represented as being the same. It appears that disturbances of motion continuity, and topological changes, are especially likely to break the representation of sameness (von Marle and Scholl 2003; Zhou et al. 2010).

MOT-type tasks have several features by virtue of which they are useful for investigating the transitivity of visual sameness. First, to successfully resolve an MOT task, the visual system has to track the sameness of each target through a period of movement. An alternative strategy might be to treat all targets as a single group and check whether an object distinguished at the end of a trial belongs to this group without individually tracking each target (Yantis 1992). While it is true that grouping cues make tracking easier (Makovski and Jiang 2009), it is not likely to be a sufficient explanation of performance in MOT. In particular, if objects move at random and have the same properties, such as colour and shape, then for successful target grouping each has to be tracked separately. Without this, a distractor can easily be included in a group of targets, while one of the targets is excluded, as there are no additional cues that would allow participants to distinguish members of the target group (Pylyshyn 2004).

Second, splitting-like situations are quite common in MOT-type tasks. When two objects move close to one another, there are two subsequent moments: T<sub>1</sub> when objects A and B are present, and T<sub>2</sub> when there are objects C and D (Fig. 8). All the objects have the same colour, size, and shape. In addition, each of the objects at T<sub>1</sub> is spatiotemporally continuous with the objects at T<sub>2</sub>, as the temporal distance between these moments is short, and the locations of A and B are close to the locations of C and D. Therefore both C and D seem to be candidates for being the same as A, as well as for being the same as B.

These splitting-like situations are stronger than those connected with the standard streaming/bouncing stimuli (Sect. 2), because when several objects are tracked simultaneously, the visual system is not able to precisely predict the future positions of

perceived items (Iordanescu et al. 2009; Keane and Pylyshyn 2006). This limitation suggests that in many splitting-like situations occurring in MOT, visual mechanisms would not be able to resolve the ambiguity by establishing sameness between earlier and later objects relying on the predicted localisation.

In works concerning MOT studies, it is universally accepted that many identification errors occur when two objects are close to each other, as higher frequencies of such situations lead to worse performance (Intriligator and Cavanagh 2001; Alvarez and Franconeri 2007; Tombu and Seiffert 2011). This additionally shows that when objects in MOT are in proximity, such that splitting-like situations occur, something interesting happens with the object's sameness.

Third, the design of MOT studies allows us to distinguish visual sameness from spatiotemporal continuity. As stated earlier, when a splitting-like situation is investigated by measuring OSPB, it is not clear whether this effect is present due to the occurrence of sameness, or whether it is present due to appropriate spatial relations. In contrast, within the MOT paradigm, responses given by participants are directly about the sameness of objects.

Finally, and in contrast to investigations using streaming/bouncing stimuli, it seems plausible that results obtained in MOT studies are connected to the functioning of perceptual mechanisms rather than higher-order reasoning. Simultaneous tracking of multiple objects consumes a significant amount of cognitive resources, which cannot be then used to form high-level judgments concerning sameness. Accordingly, MOT participants are not able to reflect on, and modify, represented sameness relations. The majority of psychological models of MOT confirm this by stating that target-tracking happens as a result of relatively low-level processes (Scholl 2001; Oksama and Hyönä 2004).

In further subsections, I analyse three MOT-related results that are relevant to questions concerning the transitivity of sameness: asymmetry between target/distractor and target/target errors (Sect. 4.1), measurements of Contralateral Delay Activity (Sect. 4.2), and proportions of two types of target/distractor errors (Sect. 4.3).

#### 4.1 Asymmetry of errors

An important version of the MOT task is Multiple Identity Tracking (MIT). In this type of study objects do not have the same features because (a) at the beginning of a trial, targets are marked with individual numbers that later disappear when the objects start to move (Pylyshyn 2004), or (b) all objects, whether targets or distractors, have different features during the whole trial (Oksama and Hyönä 2008). When the latter option is applied, all objects hide behind uniform occluders at the moment the movement stops. In MIT, the participants' task is not to decide whether a given object is a target or a distractor, but, depending on the experimental design, to (a) assign to each target the same number that it carried at the beginning of a trial, or (b) decide the location of an object (i.e., behind which occluder) with a given combination of features.

The MIT design allows two types of mistakes to be identified: target/distractor errors, which consist of mistaking a target for a distractor, and target/target errors in

which the identities of two targets are confused. Regardless the specific design that is applied, target/target errors are significantly more frequent than target/distractor errors (Horowitz et al. 2007; Pinto et al. 2010). Target/target errors usually happen when two targets move close to each other, i.e. when the splitting-like situations are likely to occur (Pylyshyn 2004, pp. 819–820).

The above asymmetry between two types of error has an important implication for investigations concerning splitting-like situations. Specifically, it provides evidence against the “End of sameness” solution. In a splitting-like situation involving targets, four objects should be considered (Fig. 8): object A, B at a moment  $T_1$ , object C, and object D at a subsequent moment  $T_2$ . If such a splitting-like situation breaks sameness, then objects at  $T_1$  should not be the same as any objects at  $T_2$ . In this case, two options are available: objects C and D may be treated as distractors, or they may be new targets, which are not the same as A or B.

The first option, according to which objects at  $T_2$  are distractors, is unlikely given the asymmetry of errors. If targets after a splitting-like situation are no longer treated as targets, but as distractors, then they are no longer tracked and would be easily confused with other distractors. In such a case, every splitting-like situation, and so the majority of target/target errors, will be followed by target/distractor errors. As such, the frequency of target/distractor errors should not be significantly lower than the frequency of target/target errors. However, the obtained results demonstrate that in MIT tasks, often two targets are mistaken for each other, yet the mistaken targets are still successfully discerned from distractors.

According to the second option, splitting-like situations between targets result in the representation of new targets. This alternative is not very plausible given the results of MIT studies. In designs where numbers are attributed to targets at the beginning of a trial, it was observed that target/target errors do not happen at random (Pylyshyn 2004, pp. 815–819). Instead, when two targets are close to each other, let's say target X with the number 1 attributed to it and target Y with the number 2, their identities can “swap”, and at the end of the task participants assign 2 to X and 1 to Y. Hence, not every pattern of error is possible at every trial. To demonstrate this, let's consider a trial with four targets with numbers assigned: X-1, Y-2, V-3, and Z-4. Let's assume that exactly two splitting-like situations occurred in this trial: between X and Y, and between V and Z. As a result, exactly two errors are possible: number 2 can be attributed to X and number 1 to Y, and number 3 can be attributed to V and number 4 to Z. However, if splitting-like situations lead to a representation of new targets, this non-random pattern of errors should not happen, since the observed target/target errors should not depend on splitting-like situations. Two new targets “produced” in a splitting-like situation would not be represented as having any numbers assigned to them, since they are not treated as being the same as objects that obtained numbers at the beginning of a trial. After the two splitting-like situations described above, all targets would be new ones without numbers, so all mistakes in target-number attribution should be equally probable, since participants' responses result from guessing. However, such random patterns are not likely, given the observed dependency of errors on targets' proximity.

On the other hand, the phenomenon of asymmetry of errors is consistent with “Unique sameness”. If each object at  $T_1$  is represented as being the same as exactly one object at  $T_2$ , and each object at  $T_1$  is also represented as being the same as a

different object at  $T_2$ , then we can plausibly explain why there are more target/target errors than target/distractor ones. When targets are close to each other they may swap their identities in a random fashion, which can easily lead to target/target errors, but objects resulting from such a splitting-like event are still regarded as targets, therefore they are not easily confused with distractors. Furthermore, “sameness swapping” is consistent with observed non-random patterns of target/target errors (Pylyshyn 2004, p. 818).

The intransitive “Double sameness” is also coherent with the asymmetry of errors. In this case each object at  $T_1$  is represented as the same as both objects at  $T_2$ . After a splitting-like situation, each object has a double identity, inherited from the objects that entered this situation. This can lead to a target/target error because a participant, at the end of a task, is forced to arbitrarily disambiguate the situation by attributing a single identity to a single object. However, target/target errors will be more frequent than target/distractor ones since after a splitting-like situation objects are still represented as targets. In addition, “Double sameness” preserves the dependency of possible errors on occurrences of splitting-like situations. Using the previous example, if there are targets with attributed numbers, X-1, Y-2, V-3, Z-4, and two splitting-like situations occur: X/Y and V/Z, then two errors are possible: (a) X can be given number 2 and Y number 1, and (b) V can be given number 4 and Z number 3. According to “Double sameness”, after these splitting-like situations, all targets have double identities, X-1/2, Y-1/2, V-3/4, Z-3/4, and attempts to attribute one identity to one target can lead to exactly two errors—(a) and (b)—characterised above.

Overall, the phenomenon of asymmetry of errors provides another reason for rejecting “End of sameness”. However, it seems equally coherent with “Unique sameness” and “Double sameness”.

## 4.2 Contralateral delay activity

Authors of MOT-related works often claim that errors in the re-identification of targets have two main sources. First, they result from proximity between objects, here interpreted as leading to splitting-like situations, because the number of errors is higher when objects meet more often (Tombu and Seiffert 2011; Bae and Flombaum 2012). Second, a tracked target may be lost, even without an interaction with another object, if insufficient perceptual resources are assigned to it. Such mistakes occur more often if, *inter alia*, the velocity of objects is higher (Feria 2013).

The occurrences of target/distractor errors of the two above variants, connected with proximity or velocity, were investigated by conducting electrophysiological measurements of Contralateral Delay Activity (CDA) (Drew et al. 2013). The investigations revealed that CDA is higher when the number of tracked targets is larger (Drew and Vogel 2008). This correlation allows us to use CDA measurements to evaluate how splitting-like situations are resolved (see Skrzypulec 2016 for a preliminary exposition of this idea).

When a target and a distractor are in proximity, a splitting-like situation has the following form: at  $T_1$  there are two objects, one target and one distractor, and at a subsequent moment  $T_2$  there are also two objects. In this case, the target from  $T_1$  may

be represented as being the same as exactly one object at  $T_2$  (“Unique sameness”). According to “End of sameness”, sameness is not represented between any objects at  $T_1$  and  $T_2$ , and (A) objects at  $T_2$  may be two distractors or (B) one of them may be a new target. Finally, if such splitting-like situations are resolved, as described in “Double sameness”, the target at  $T_1$  is represented as being the same as both objects at  $T_2$ .

If splitting-like situations between targets and distractors are resolved in accordance with “Unique sameness” or (B) variant of “End of sameness”, the CDA should remain constant, because these two solutions do not change the number of targets, so the level of CDA should not change. By contrast, if variant (A) of “End of sameness” is true, then CDA should be lower after splitting-like situations. According to this solution, there is one target at  $T_1$  and no targets at  $T_2$ , so the overall number of targets is lower after a splitting-like situation. But rising CDA is consistent with “Double sameness”. In this case, in a splitting-like situation between a target and a distractor, there is one target at  $T_1$  but two targets at  $T_2$ . As a result, the number of targets is larger, leading to higher CDA.

In one experiment, Drew and colleagues (Drew et al. 2013, pp. 215–216) manipulated the number of distractors. A rising number of distractors causes a higher number of target/distractor errors due to more cases of close encounters between objects. Some of these errors are likely to arise from splitting-like situations. However, other factors influencing the rate of errors, like the velocity of objects, has not been changed. The results showed that while higher distractors numbers led to more frequent target/distractor errors, the CDA did not change, suggesting that there was no change in the number of objects represented as targets (Drew et al. 2013, pp. 216–217). This is consistent with “Unique sameness” and the (B) variant of “End of sameness”, but is inconsistent with variant (A) of “End of sameness” and with “Double sameness”.

The above result corroborates the hypothesis that visual sameness is transitive by being inconsistent with “Double sameness”. However, it does not allow us to decide whether splitting-like situations break the sameness (variant (B) of “End of sameness”) or not (“Unique sameness”).

### 4.3 Proportions of target/distractor errors

When a participant makes a target/distractor error during an MOT task, it can mean one of two things. First, that the question at the end of a trial was about a target, but the participant’s response was “distractor”. I will call such mistakes *target*  $\gg$  *distractor* errors. Second, a question might be about a distractor, but the answer given was “target”. These are *distractor*  $\gg$  *target* errors. Below, I demonstrate how, by investigating which of these two types of mistakes is more frequent, we may obtain another clue regarding the transitivity of visual sameness.

As stated earlier (Sect. 4.2), there are two main sources of mistakes in MOT tasks: dropping tracked targets (e.g. due to velocity) and splitting-like situations, when objects move close to each other. If a target is dropped during tracking, two things can happen (Pylyshyn et al. 1994; Zelinsky and Neider 2008): (a) one of the objects that is in proximity to the last recorded location of a dropped target is recovered as



this target; (b) there are no attempts to recover a dropped target, and it is no longer distinguished from distractors. After a recovery attempt (option (a)), one of the distractors may be represented as a target, with the actual dropped target represented as a distractor. If this is the main source of target/distractor errors, then the number of *target*»*distractor* and *distractor*»*target* errors should be similar, because after an unsuccessful recovery there is one target that is wrongly represented as a distractor, and one distractor that is treated as target. However, if there is no recovery attempt (option (b)), *target*»*distractor* errors are more likely. In these cases, when a target is dropped it becomes a distractor, so one target is incorrectly represented as a distractor, but no distractor is treated as target. Because it is likely that both options (a) and (b) occur in MOT tasks, a higher number of *target*»*distractor* errors should be observed, if droppings were the only sources of target/distractor errors.

It can now be asked how this pattern changes if target/distractor errors also result from splitting-like situations. Resolutions of splitting-like situations along “Unique sameness” do not make either of the two types of target/distractor mistakes more probable. In this case, a target at an earlier moment  $T_1$  is represented as being the same as exactly one object at  $T_2$ . If this leads to an error, then at  $T_2$  there is a distractor represented as a target, and a target represented as a distractor. The effect of the (B) variant of “End of sameness” will be the same. Here, a target at  $T_1$  is not the same as any object at  $T_2$ , but one of them is a new target. However, if variant (A) of “End of sameness” is right, then the *target*»*distractor* errors should be more frequent than *distractor*»*target* ones, because at  $T_2$  there are two distractors, so in a whole display there is a target that is wrongly represented as a distractor, but all distractors are represented correctly.

Each of these three solutions, “Unique sameness”, and two variants of “End of sameness”, gives the same result when combined with the effect of target dropping: *target*»*distractor* errors should be more frequent. “Unique sameness” and the (B) variant of “End of sameness” predict that both types of errors should be equal in number, but the influence of target dropping leads to a slightly larger proportion of *target*»*distractor* mistakes. The (A) variant of “End of sameness” favours *target*»*distractor* errors in the same way as target dropping.

The results should differ according to “Double sameness”. In this case, both objects at  $T_2$  are targets represented as being the same as a single target at  $T_1$ , so there is a distractor that is incorrectly interpreted as a target, but there is no target incorrectly identified as a distractor.<sup>4</sup> Hence, *distractor*»*target* mistakes should be more common than *target*»*distractor* ones. This effect should weaken the influence of target droppings, which favour *target*»*distractor* errors. Hence, if *target*»*distractor* errors dominate over *distractor*»*target* ones, this counts against the intransitive “Double sameness”.

In most of the MOT-related papers, variances in frequency between two types of target/distractor errors are not reported. However, when they are reported, they show the dominance of *target*»*distractor* errors (Sears and Pylyshyn 2000, p. 11; Oksama

<sup>4</sup> One may argue that each object at  $T_2$  is both a target and a distractor. However, it is not clear whether sameness of distractors is represented during MOT (participants are not able to re-identify distractors), and the task itself is partially about neglecting distractors. Hence it seems that objects at  $T_2$  are treated primarily as targets.

and Hyönä 2004, p. 648). The study by Oksama and Hyönä seems to be of particular relevance, as it includes a much higher number of participants (about 200) than is usual in cognitive psychology. This result is problematic for the “Double sameness” solution, thereby corroborating the hypothesis that visual sameness is transitive. However, it does not allow us to decide whether splitting-like situations break visual sameness (“End of sameness”) or not (“Unique sameness”).

## 5 Unique sameness and primitive sameness

The results analysed in the previous sections do not allow us to infer a final solution to the question of the transitivity of visual sameness. However, they show that some answers are more probable than others, according to the current state of research. In particular, “Unique sameness” is consistent with all five considered results: (a) peoples’ judgments regarding streaming/bouncing stimuli, (b) the presence of OSPB in cases of division, (c) the asymmetry of target/target and target/distractor errors, (d) measurements of CDA, and (e) the frequency of *target*≫*distractor* and *distractor*≫*target* mistakes. The intransitive “Double sameness” is consistent with two of these results: OSPB in cases of division, and asymmetry of target/target and target/distractor errors. Similarly, “End of sameness”, in one of its two versions, is also consistent with two of them: measurements of CDA, and the frequency of *target*≫*distractor* and *distractor*≫*target* mistakes.

Given the above results, the most plausible model of visual splitting-like situations is that they are resolved as described in “Unique sameness”. This means that in a splitting-like situation, an object at an earlier moment is represented as being exactly the same as one object at a later moment. Such a result corroborates the hypothesis that the visual sameness is transitive. However, it should be noted that, technically, a single counterexample is sufficient to show that visual sameness is not transitive, and it is possible that further investigations into OSPB in division cases will refute the “Unique sameness” interpretation. Such an outcome would suggest that, while in most splitting-like cases visual sameness behaves like transitive relation, it is not the case in special patterns, when one object divides into two (see Sect. 6.1).

The transitive interpretation described above has an interesting consequence for our understanding of the content of visual states. According to an important tradition in philosophy of perception (e.g., Russell 1956, pp. 337–339), vision represents objects as combinations of features and places. The contents of many visual states, connected with important psychological phenomena, may be described in this way. For example, when a figure is distinguished from ground, it is represented that there is an object composed of, *inter alia*, a certain shape and colour, as well as properties describing borders. Such an object may be represented as standing in spatial relation to other objects, consequently it can be represented as a part of a more complex object, or as a member of a perceptual group of similar objects (Palmer and Rock 1994).

However, if the “Unique sameness” solution to splitting-like situations is the correct one, the above picture is incomplete. To demonstrate this, let’s consider a splitting-like situation with a single object A at  $T_1$  and two objects, B and C, at  $T_2$ . According to “Unique sameness”, the object A is represented as being the same as one of the objects

at  $T_2$ . Let's assume that it is the object B. In this case, a legitimate question would be: for what reason was B, rather than C, chosen to be the same as A? The earlier discussion showed that it is plausible to claim that there are visual splitting-like situations, for example occurring in MOT, where there is no bias, arising from similarity of shape and colour, spatiotemporal continuity, or properties concerning predicted future location, that determine which of the later objects should be represented as being the same as an earlier object.<sup>5</sup>

One may propose that in our splitting-like situation the object B was chosen because it possesses a temporal property such as “being in place P at  $T_1$ ”, where P is the location of the object A at  $T_1$ , while the object C lacks this property. However, it seems plausible, that to represent that an object B presented at a moment  $T_2$  had property F at an earlier moment  $T_1$ , it is necessary to recognise that (a) there were some objects at  $T_1$ , (b) that one of them had F, (c) that one of them is the same as B, and finally (d) that the object F at  $T_1$  is the same as B. So representing temporal properties already presupposes representing sameness, and outcomes of splitting-like situations cannot be explained by differences in temporal properties.

The presence of such strong splitting-like situations, resolved in accordance with “Unique sameness”, shows that we cannot explain in virtue of what one of the later objects was chosen to be the same as an earlier object by referring to the usual elements of visual content (e.g. colours, shapes, or spatial relations). Using a term taken from analytic metaphysics, it can be stated that visual object A, in our earlier example, is represented as “primitively the same”<sup>6</sup> as visual object B, since representation of their sameness does not supervene on their represented intrinsic or relational qualitative properties. The concept of “primitive sameness” has played an important role in metaphysical discussions concerning modalities (Adams 1979), but was not applied in considerations of visual content.

To somehow ground primitive sameness in the structure of objects, a special element is postulated, characterised as “thisness”, i.e. a non-qualitative property like “being the object A” (see Ladyman 2007 for a contemporary descriptions). Apart from various characterisations of this special element, it must possess two formal features. First, the identity of “thisness” should be a part of necessary and sufficient conditions of the sameness of objects. Second, contributing to the sameness of objects should be its sole function. Hence, “thisness” is different from more usual properties of objects, which provide qualitative characteristics (like “being red”).

It should be noted that a reference to “thisness” and “primitive sameness” cannot be simply omitted by claiming that in splitting-like situations, resolved in accordance with “Unique sameness”, sameness is established arbitrarily or randomly (similar reasoning has been presented earlier by Skrzypulec 2018). In fact, a reference to “thisness” is

<sup>5</sup> As noted in Sects. 1 and 4, while objects in the considered experiments have distinct spatial trajectories, the empirical results suggest that information about such spatial differences is poorly represented (Intriligator and Cavanagh 2001; Alvarez and Franconeri 2007; Tombu and Seiffert 2011). It is particularly the case when, as in MOT experiments, several objects are tracked simultaneously. As a result, even if objects engaged in a splitting-like situation may have differing trajectory-related properties, it is likely that these differences are not visually represented.

<sup>6</sup> The usual philosophical term is “primitive identity” but as my investigations do not prove that visual sameness is the identity relation (because being transitive is a necessary but not sufficient condition for being the identity relation), I use the term “primitive sameness”.

needed to properly formulate sameness conditions of objects that are represented as being the same due to a random process. Let's once again consider a splitting-like situation with a single object A at  $T_1$  and two objects, B and C, at  $T_2$ . In addition, let's assume that the sameness relation is established between A and B by a random process. In such a splitting-like situation, it is represented that an object A exists at moment  $T_1$  and that the same object A also exists at a subsequent moment  $T_2$  (as it is represented as numerically the same as an object B). At  $T_2$  an object C is also represented as being different from A. Because sameness is established randomly, necessary and sufficient conditions of objects' sameness cannot be formulated by referring solely to spatial relations and qualitative properties, since they do not determine which of the objects at  $T_2$  is represented as the same as A at  $T_1$ . It seems that there is no other way to formulate sameness conditions than by referring simply to the individualising characteristics, expressed by phrases like "being the object  $\varphi$ " (e.g., A or B), which characterise objects as being the same or different. Such an individualising characteristic is a form of "thisness", which grounds "primitive sameness".

According to the above considerations, if strong splitting-like situations are resolved along "Unique sameness", then in some cases the visual system establishes relations of visual sameness in a way that is grasped by the concept of "primitive sameness". This is important for the proper characterisation of visual content, as vision represents objects not only as combinations of usual features and places, but also as constituted by an additional individualising element, with the formal characteristics of "thisness". Hence, adopting "Unique sameness" as the most probable solution to splitting-like situations, not only provides an argument for the transitivity of visual sameness, but also suggests that the concepts of "thisness" and "primitive sameness" are required in order to account for the way in which objects are visually represented.

## 6 Alternative models

According to the most plausible model of visual sameness, it is a transitive relation, and there are cases in which it should be characterised as a "primitive sameness". However, relying on the current state of knowledge, it is impossible to decisively reject all alternative models of visual sameness. As such, it is worth considering the psychological results that would justify adopting a different model of the considered relation.

There are three alternative models that can be easily constructed by negating features of the "main model" described in the previous section. First, visual sameness may be characterised as intransitive, but involving cases of "primitive sameness". Second, it can be transitive, but without needing to invoke the concept of "primitive sameness". Finally, visual sameness may be treated both as intransitive and not involving "primitive sameness".

### 6.1 Intransitivity and "primitive sameness"

The first, and probably the most plausible of the above alternative models, would be justified if two claims are true: (1) there are splitting-like situations that are resolved in

accordance with “Double sameness” and (2) there are splitting-like situations resolved by characterising them as “Unique sameness”, where patterns of sameness between earlier and later objects cannot be explained in terms of differences in qualitative properties. The truth of (1) entails that the relation is intransitive, while the truth of (2) justifies occurrences of “primitive sameness”.

The plausibility of (2) was defended in the previous section, where the main model was characterised (Sect. 5). However, the considered results do not provide a clear case of a splitting-like situation that is resolved in accordance with “Double sameness” as described in (1). Probably the best candidates are situations of division investigated using the OSPB paradigm (Sect. 3). In a case where one object divides into two, the “Unique sameness” interpretation assumes that one of the later objects is the same as the earlier one, and the second later object is a completely new entity. It might be the case that the visual system rejects such an unexpected appearance of a new object, and interprets a division situation in accordance with “Double sameness”, while other splitting-like situations (e.g. those occurring in MOT studies) are resolved along “Unique sameness”.

As I stated earlier, the results of OSPB investigations are coherent both with “Unique sameness” and “Double sameness”, and researchers have chosen to remain agnostic about the correct interpretation. Other studies involving splitting-like situations do not favour “Double sameness”, and it is not clear whether OSPB is really connected with sameness rather than merely spatial relations. However, if further investigations prove that OSPB concerns sameness, and the interpretation of division cases in terms of “Unique sameness” has to be rejected, then the main model of visual sameness should be abandoned in favour of an intransitive one.

## 6.2 Transitivity without “primitive sameness”

According to the second of the alternative models, visual sameness is transitive, but it never has the characteristic of “primitive sameness”. This model would be justified if both claims (1) and (2) are false, i.e. no splitting-like situations are resolved according to “Double sameness” and in all splitting-like situations resolved in accordance with “Unique sameness” patterns of sameness can be explained in terms of differences in qualitative properties.

The falsity of (1) is also needed for the adequacy of the main model, and in the previous sections it was argued that rejection of (1) seems plausible. However, it is more difficult to reject (2). This would be justified if all splitting-like situations break sameness as described by “End of sameness”; but this is very unlikely given the analysed data.

Another option is to argue that even if there are splitting-like situations that can be resolved in accordance with “Unique sameness”, there is always a difference in qualitative properties that determines the patterns of visual sameness. One idea is to use the fact that in MOT studies, when several objects are tracked, objects’ properties are represented imprecisely, due to the division of cognitive resources between targets (Feria 2012; Howard and Holcombe 2008). Because of such limits, it may be the case that when a splitting-like situation happens, spatiotemporal continuity is rep-

resented only between one earlier object and one later object, without the possibility of simultaneously representing continuity between multiple objects. Hence, splitting-like situations occurring in MOT would be rather weak, since there would always be a qualitative factor determining the pattern of visual sameness.

If the above scenario were proven to occur, this would constitute an argument against the main model, and in favour of an alternative model, that rejects the occurrences of “primitive sameness”.

### 6.3 Intransitivity without “primitive sameness”

The last of the alternative models would be justified if (1) is true, but (2) is false: there are splitting-like situations resolved in accordance with “Double sameness”, but there are no splitting-like situations resolved as described in “Unique sameness”, where patterns of sameness cannot be explained by reference to qualitative properties.

Results that would suggest adopting this model are a combination of those justifying the two previous alternative models. First, (1) would be justified if OSPB occurs as a result of representing visual sameness, and division cases are resolved as characterised by “Double sameness”. Second, (2) can be rejected by successfully reinterpreting strong splitting-like situations which occur, *inter alia*, during MOT.

The above considerations show that, while the main model of visual sameness—according to which sameness is transitive, and there are cases of “primitive sameness”—is the most plausible (Sect. 5), there are possible results that would lead to its abandonment in favour of one of the alternative models. In addition, of the alternative models, the most serious candidate for replacing the main model can be estimated. Among the three alternative models, the one characterising visual sameness as intransitive, but involving “primitive sameness” is the easiest to argue for (Sect. 6.1). All that is needed for its accuracy is the presence of any splitting-like situations that are resolved in accordance with “Double sameness”.

## 7 Conclusions

The goal of the paper was to investigate whether the relation of visual numeric sameness should be characterised as transitive, according to the current state of empirical knowledge. Five experimental results regarding the occurrences of visual splitting-like situations have been analysed. I argued that the transitive interpretation of visual sameness is the most plausible. Further, the considered data suggest that some occurrences of visual sameness should be classified as representations of “primitive sameness”.

**Acknowledgements** The author would like to thank participants of CLMPS 2015 conference, participants of cognitive science seminar at Jagiellonian University, participants of 10th Polish Congress of Cognitive Science and three anonymous reviewers for their comments concerning the earlier versions of the paper. The work was supported by the National Science Center (Poland) Grant 2014/12/T/HS1/00249.

## Compliance with ethical standards

**Conflict of interest** The author declares that he has no conflict of interest.

**Ethical standards** I comply with the ethical standards of „Synthese“.

**Informed consent and animal welfare** The research did not involve human participants or animals.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Adams, R. M. (1979). Primitive thisness and primitive identity. *The Journal of Philosophy*, 76(1), 5–26.
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, 7(13), 14.
- Atsma, J., Koning, A., & van Lier, R. (2012). Multiple object tracking: Anticipatory attention doesn't "bounce". *Journal of Vision*, 12(13), 1–11.
- Bae, G. Y., & Flombaum, J. I. (2012). Close encounters of the distracting kind: Identifying the cause of visual tracking errors. *Attention Perception and Psychophysics*, 74, 703–715.
- Cacchione, T., & Call, J. (2011). Do gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) fail to represent objects in the context of cohesion violations? *Cognition*, 116, 193–203.
- Caplovitz, G. P., Shapiro, A. G., & Stroud, S. (2011). The maintenance and disambiguation of object representations depend upon feature contrast within and between objects. *Journal of Vision*, 11(14), 1–14.
- Chisholm, R. (1969). The Loose and Popular and the Strict and Philosophical Senses of Identity. In N. Care & H. Grimm (Eds.), *Perception and identity* (pp. 82–106). Cleveland: Case Western Reserve University Press.
- Chuard, P. (2007). Indiscriminable shades and demonstrative concepts. *Australasian Journal of Philosophy*, 85(2), 277–306.
- Dawson, M. R. W. (1990). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review*, 98(4), 569–603.
- Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition*, 126, 213–223.
- Drew, T., & Vogel, E. K. (2008). Neural measures of individual differences in selecting and tracking multiple moving objects. *Journal of Neuroscience*, 28(16), 4183–4191.
- Feria, C. S. (2012). The effects of distractors in multiple object tracking are modulated by the similarity of distractors and target features. *Perception*, 41, 287–304.
- Feria, C. S. (2013). Speed has an effect on multiple-object tracking independently of the numbers of close encounters between targets and distractors. *Attention, Perception, and Psychophysics*, 75, 53–67.
- Graff, D. (2001). Phenomenal continua and the sortites. *Mind*, 110, 905–935.
- Hardin, C. L. (1988). Phenomenal color and sortites. *Noûs*, 22, 213–234.
- Hollingworth, A., & Franconeri, S. L. (2009). Object correspondence across brief occlusion is established on the basis of both spatiotemporal and surface feature cues. *Cognition*, 113, 150–166.
- Hood, B. M., & Bloom, P. (2008). Children prefer certain individuals over perfect duplicates. *Cognition*, 106, 455–462.
- Horowitz, T. S., Klieger, S. B., Fencsik, D. E., Yong, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Perception and Psychophysics*, 69(2), 172–184.
- Howard, C. J., & Holcombe, A. O. (2008). Tracking the changing features of multiple objects: Progressively poorer perceptual precision and progressively greater perceptual lag. *Vision Research*, 48, 1164–1180.
- Howe, P. D. L., & Holcombe, A. O. (2012). The effects of visual distinctiveness on multiple object tracking performance. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2012.00307>.

- Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology*, *43*, 171–216.
- Iordanescu, L., Grabovecky, N., & Suzuki, S. (2009). Demand-based dynamic distribution of attention and monitoring of relocation during multiple-object tracking. *Journal of Vision*. <https://doi.org/10.1167/9.4.1>.
- Julesz, B. (1995). *Dialogues on perception*. Cambridge, MA: MIT Press.
- Kahneman, D., Treisman, A. M., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.
- Keane, B., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*, 346–368.
- Ladyman, J. (2007). On the identity and diversity of objects in structures. *Proceedings of the Aristotelian Society, Supplementary Volumes*, *81*, 23–43.
- Lewis, D. (1986). *On the plurality of worlds*. Oxford: Blackwell.
- Makovski, T., & Jiang, Y. V. (2009). The role of visual working memory in attentive tracking of unique objects. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1687–1697.
- Mills, E. (2002). Fallibility and the phenomenal sorites. *Noûs*, *36*(3), 384–407.
- Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin and Review*, *14*(6), 1199–1204.
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2004). Divide and Conquer. How object files adapt when a persisting object splits into two. *Psychological Science*, *15*(6), 420–425.
- Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition*, *96*, 67–92.
- Moore, C. M., Stephens, T., & Hein, E. (2010). Features, as well as space and time, guide object persistence. *Psychonomic Bulletin and Review*, *17*, 731–736.
- Natsoulas, T. (1992). The tunnel effect, Gibson's perception theory, and reflective seeing. *Psychological Science*, *54*, 160–174.
- Nudds, M. (2014). Auditory appearances. *Ratio*, *27*, 462–482.
- O'Callaghan, C. (2014). Auditory perception. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), E. N. Zalta (ed.). <https://plato.stanford.edu/archives/win2016/entries/perception-auditory/>.
- Odic, D., Roth, O., & Flombaum, J. I. (2012). The relationship between apparent motion and object files. *Visual Cognition*, *20*(9), 1052–1081.
- Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*(5), 631–671.
- Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology*, *56*, 237–283.
- Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin and Review*, *1*(1), 29–55.
- Parfit, D. (1971). Personal identity. *The Philosophical Review*, *80*(1), 3–27.
- Pinto, Y., Howe, P. D. L., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of Vision*, *10*(10), 1–15.
- Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*, 127–158.
- Pylyshyn, Z. W. (2004). Some puzzling finding in multiple object tracking: I. Tracking without keeping track of object identities. *Visual Cognition*, *11*(7), 801–822.
- Pylyshyn, Z. W. (2007). *Things and places*. Cambridge, MA: MIT Press.
- Pylyshyn, Z. W., Burkell, J., Fisher, B., Sears, C., Schmidt, W., & Trick, L. (1994). Multiple parallel access in visual attention. *Canadian Journal of Experimental Psychology*, *48*, 260–283.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.
- Raffman, D. (2000). Is perceptual indiscriminability nontransitive? *Philosophical Topics*, *28*(1), 153–175.
- Rips, J. (2011). Split identity: Intransitive judgments of the identity of objects. *Cognition*, *119*, 356–373.
- Russell, B. (1956). *An inquiry into meaning and truth*. London: George Allen and Unwin.
- Scholl, B. J. (2001). Objects and attention: The state of art. *Cognition*, *80*, 1–46.
- Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind and Language*, *22*(5), 563–591.
- Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple-object tracking. *Cognition*, *80*(1/2), 159–171.



- Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Psychology*, *54*(1), 1–14.
- Sekuler, A., & Sekuler, R. (1999). Collisions between moving visual targets: What controls alternative ways of seeing an ambiguous display? *Perception*, *28*, 415–432.
- Skrzypulec, B. (2016). Commentary: Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2016.00372>.
- Skrzypulec, B. (2018). Thisness and visual objects. *Review of Philosophy and Psychology*, *9*(1), 17–32.
- Tombu, M., & Seiffert, A. E. (2011). Tracking planets and moons: Mechanism of object tracking revealed by a new paradigm. *Attention, Perception and Psychophysics*, *73*, 738–750.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- von Marle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14*(5), 498–504.
- Williamson, T. (2013). *Identity and discrimination*. Chichester: Wiley-Blackwell.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, *30*, 111–153.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.
- Zelinsky, G., & Neider, M. (2008). An eye movement analysis of multiple object tracking in a realistic environment. *Visual Cognition*, *16*, 553–566.
- Zhou, K., Luo, H., Zhou, T., Zhou, Y., & Chen, L. (2010). Topological change disturbs object continuity in attentive tracking. *PNAS*, *107*(50), 21920–21924.