Liz Irvine, 6575 words

<div align="center">Memory Images</div>

Introduction

The aim of this chapter is to assess arguments, based on research into visual short term memory, that the contents of consciousness are more rich and detailed than we can evidence in reports. Several authors claim that the rich and possibly non-conceptual contents of iconic memory contribute to the contents of consciousness, but much of this is not processed into working memory so is soon 'forgotten' (Block, 2011, 2007; Dretske, 2007, 2006, 2004; Fodor, 2008; Tye, 2009, 2006). This short-lived, non-attended and unreported conscious content is argued to exist on a 'phenomenal' (Block) or 'object' (Dretske) level of awareness. Content that is attended and reported on is cognitively accessed (Block), or experienced on a 'fact' level (Dretske) as well.

This reflects Block's well known distinction between phenomenal and access consciousness. Phenomenal consciousness refers to content that is experienced, and access consciousness refers to whatever content is made available to consumer systems, including those systems that generate (phenomenological) reports. Block uses evidence about short term visual memory to lend support to the idea that phenomenal consciousness can exist without access: we can experience things we cannot act or report on. This particular claim is situated within a biological account of consciousness, where converging lines of empirical evidence are used to argue that some contents are present in consciousness, even if the subjects does not (perhaps cannot) report on that content specifically (Block, 2011, 2007; Lamme, 2006).

This claim provides a challenge to currently popular Global Workspace theories of consciousness (Dehaene et al., 2011) which suggest that conscious content is limited to the contents of working memory. The basic idea behind these accounts is that the contents of consciousness are fairly limited at any one time, but change as shifts of attention populate working memory with new content. This position in turn often stems from a mix of functionalist and operationalist perspectives on consciousness (e.g. Cohen and Dennett, 2011). If the contents of consciousness must be accessible to the subject (and so to experimenters via report) to serve any function, then if some content appears to be currently unaccessed or unreportable, then there is no reason to think of it is as being part of the contents of consciousness. This is a straightforward denial of the idea that phenomenal consciousness can exist without accessibility. So, since the contents of working memory are those contents that are currently accessed and reportable, this means that the contents of consciousness are tied to the contents of working memory (non-overflow accounts).

To understand this debate, this chapter focuses on evaluating the two types of short-term visual memory usually invoked to support claims about rich/overflowing phenomenal content. The first is iconic memory, which lasts less than a second, stores reasonably large amounts of visual detail, and components of which support the existence of visual after images (the 'memory images' of the title). The second is termed 'fragile visual short term memory', and is somewhere between iconic memory and working memory in both duration and capacity. The contents of these types of memory are richer than that of working memory, and are argued to form part of the contents of phenomenal consciousness, so providing short-lived but rich visual experiences.

Several broad themes emerge from discussions on how to interpret this experimental work. One recurring issue is how to conceive of the relationship between information processing, the contents of memory, subjective reports, and phenomenology. If there is a 'mis-match' between the format of processed or stored information, and what seems to be evidenced in phenomenology or in subjective reports, then it is possible to challenge simple claims that this content is present in consciousness. Rich/overflow and sparse/non-overflow accounts treat these mis-matches relationships very differently, which in turn impacts on their plausibility.

In addition, later sections raise methodological and conceptual questions about how to identify the capacity and contents of working memory. This potentially challenges *any* strong claims (overflow and non-overflow) about the relationship between working memory and consciousness. This is because if we are fundamentally unsure about what is in working memory, and what format it takes, (and how attention comes into it), then claims about the relationships between conscious content, working memory and attention become difficult to assess, both philosophically and empirically.

Arguments that iconic memory contributes content to phenomenal consciousness are reviewed first, followed by two alternative explanations of the phenomenon, and a discussion of the 'mis-match' problem. This is followed by a review of newer claims about rich conscious content based on fragile visual short-term memory, and potential consequences for both side of the debate given developing concerns about the nature of working memory. The conclusion sums up. A brief note: in these debates the devil

is in the experimental detail. The following sections are therefore quite technical in

places, but less empirically inclined readers can skim or skip as required.

Iconic Memory and Overflowing Phenomenology

The usual form of short term visual memory that is used to support claims about the

rich/overflowing content of visual experience is 'iconic memory', illustrated via

Sperling's (1960) experimental paradigm (see Figure 1).

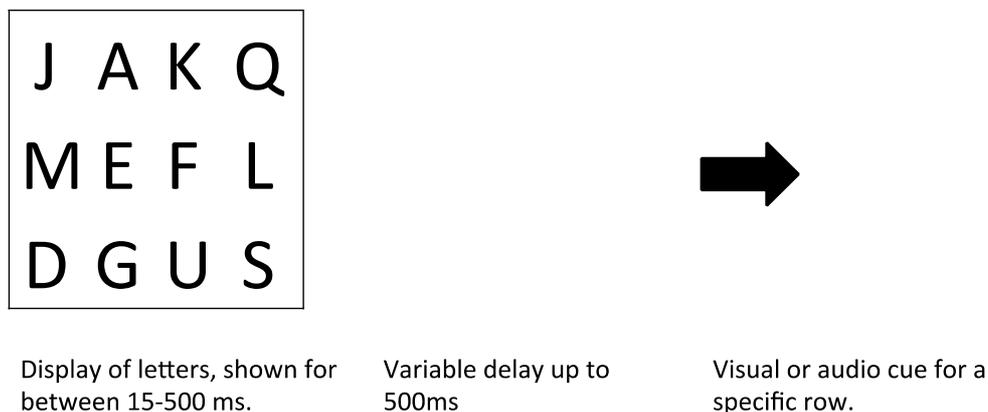| Display of letters, shown for between 15-500 ms. | Variable delay up to 500ms | Visual or audio cue for a specific row. |

Figure 1: Sperling's 'partial report' paradigm

The paradigm is fairly straightforward: subjects are shown a display of letters (usually

in a 3 by 4 grid) for a short time (15-500ms), followed by a variable delay. In the 'full

report' condition, subjects are asked to report as many letters from the display as they

can. They can typically correctly report 4.5 out of 12 letters, which is roughly what

you'd expect based on the capacity of working memory. In the 'partial report'

condition, a cue is shown to subjects just after the delay, either in the form of a visual

arrow, or a high/medium/low pitched audio tone, meant to orient attention to a

particular row in the grid. In this condition, subjects report as many letters as they can from the cued row, and on average report 3.03 letters from a 4 letter row.

The important feature of the partial report condition is that although subjects can only report 3 letters at a time, they are able to report this many letters from *any* of the rows that might be cued. In order to support this level of performance, information about 9.01 letters from the 12 letter display must therefore be available to subjects, in some form, at the time of the cue, most of which is later 'forgotten'.

Sperling's paradigm was originally seen as establishing the existence of a non-conceptual and high capacity 'iconic' visual memory store. This store appears to be short-lived and is likely pre-categorical, as only some of its content can be processed and reported at any one time. Importantly for the current discussion, subjects also report seeing 'all the letters' of the display, suggesting that the contents of this kind of iconic memory are present in consciousness.

A range of philosophers have picked up on this language of images, 'icons', and pre-categorical information, and used it to support views both about the richness of visual phenomenology and the existence of non-conceptual content. For example, Dretske (2006) notes that: ''…subjects extract [letter identity] information from what they describe as a conscious but rapidly fading image ('icon') that persists for a short time after removal of the stimulus'' (p. 175). Block (2011) says something similar in discussing the partial report condition: "…[subjects] could also report 3-4 items from any row that was cued after stimulus offset, suggesting that subjects did have a

persisting image of almost all the letters" (p. 567) (see also Block, 2007; Fodor, 2008: 189-190; Tye, 2006: 511-513).

And indeed, taking subjects' reports at face value, it does appear as though these are reasonable conclusions to draw. However, looking in more detail at contemporary work on iconic memory, and the temporal nature of consciousness, highlights alternative explanations of the Sperling paradigm, and so provide challenges to the idea of rich/overflowing visual phenomenology.

Iconic Memory and the Sperling paradigm: Where's the 'icon'?

The standard philosophical presentation of iconic memory is as a persisting visual image, from which letter identities can be read off when a cue is shown. In contrast, the contemporary scientific understanding of iconic memory suggests that it in fact refers to a range of short term memory stores. These support different visual and memory phenomena, but none of which provide what supporters of the rich/overflow view need.

One important distinction within iconic memory is between visible persistence and informational persistence. Visible persistence essentially refers to the phenomenon of visual after-images; when presented with high-contrast stimuli (e.g. bright areas on a dark background), stimulus-related activity in early visual areas can continue for up to 100ms after the stimulus has disappeared (for more see Brockmole and Wang, 2003; Di Lollo, 1980). However, what explains performance in the Sperling paradigm is not visible persistence, which is usually well over by the time the cues are shown to

subjects. Instead, performance in the Sperling paradigm is explained via the phenomenon of *informational* persistence, which is supported by a short term store of visual information. Informational persistence splits into a visible analog representation, which preserves shape and location information for 150-300ms after stimulus offset, and a nonvisual post-categorical store which preserves abstract information such as letter identity for 500 ms after stimulus offset (see e.g. Di Lollo, 1980; Irwin and Yeomans, 1986).

Given the different rates at which these different information stores decay, varying the temporal aspects of the Sperling paradigm can lead to some interesting results. For example, using a 300-500ms delay between the letter display and the cue leads to subjects making 'location errors'. They still have access to letter identities (from the post-categorical store), but not to shape and location (from the analog representation), so can correctly identify some of the letters in the display, but cannot allocate them to the correct row.

This then generates a series of problems for standard interpretations of the Sperling paradigm. First, according to the contemporary understanding of visual memory, what enables subjects to identify letters is not a persisting visual 'icon' or 'image'. The classic Sperling paradigm does not assess visible persistence, which is related to persisting visual images, only informational persistence, which is not. As Luck and Hollingworth (2008) state: "the partial-report technique [=when a row is cued] does not measure directly the visible aspect of visual sensory memory, but rather that information persists after stimulus onset'' (p. 16).

However, authors using the Sperling paradigm to argue for rich/overflowing conscious content tend to conflate visible and informational persistence. For example, in the quotation from Block above (2011), he notes that: "…[subjects] could also report 3-4 items from any row that was cued after stimulus offset, suggesting that subjects did have a *persisting image* of almost all the letters" (p. 567, italics added) (see also Block, 2007: p. 488-490, 494, 532; Tye, 2006: pp. 511–513, for similar claims). But the Sperling paradigm shows no such thing; visible persistence, which generates after-images, decays well before cues are showed to subjects, and stored (unconscious) information is all that is needed to drive subjects' performance in the task.

Second, it is also difficult to provide positive arguments for the claim that whatever persists informationally also persists phenomenologically. One issue is the existence of different memory stores, and how their contents are linked to the contents of visual consciousness. It seems to be assumed that the link is a transparent one, such that the contents of iconic memory are straightforwardly present in consciousness. However, as noted above, the different stores in informational persistence decay (and are populated) at different rates. Yet this is apparently not noticed by subjects, who only report general claims of having seen the array, or all the letters. They do not for example report phenomenological changes in where, or whether, letters are spatially located, as might be expected when letter identities are still available but spatial information is not.

If this is the case, then even if the contents of iconic memory are present in visual phenomenology, they are certainly not present in a straightforward way. No such

account of how this works has been offered, but it is necessary to support the claim that the contents of sensory memory *are* the contents of experience. Two alternative explanations are given below for what subjects experience in the Sperling paradigm that are more consistent with the contemporary science of iconic memory.

Alternative Explanations: Generic Phenomenology

Commentaries on Block (2007) abound with suggestions that subjects do not perceive all the letters in the display in detail, but that they have an experience of something like generic phenomenology, or an experience of letter-iness without an experience of specific letter shapes or identities (see also Kouider et al., 2010 on 'partial awareness'). These suggestions are certainly plausible on psychological grounds as scene gist is routinely processed quickly and efficiently, and if subjects are expecting to be presented with an array of letters, then this will be made even easier.

Indeed, the generation of reports consistent with visual richness from gist processing is a reliable phenomenon, so reliable that it can be utilized to test other features of perceptual processing. For example, Castelhano and Henderson (2008) tested how colour contributes to gist processing using a contextual bias paradigm. Here, subjects were shown a photograph of a scene for 20–250 ms followed by a 50ms mask, and then asked whether a target object was present in the photograph. For a city scene, a target object could be a fire hydrant (consistent with scene) or a tea set (inconsistent with scene). Subjects often responded that consistent items were present and that inconsistent items were not.

However, the twist in the paradigm is that *none* of the target objects were actually present in the scenes, something which subjects failed to notice. In this paradigm then, subjects generate reports consistent with having rich/overflowing experiences of specific item in the absence of item-specific processing (since here the items simply were not present). This puts pressure on accounts that take subjects' reports of having 'seen all the letters' at face value (see Irvine, 2011 for a more in depth discussion of this material).

Alternative Explanations: Postdictive Experiences

Phillips (2011a) argues for a slightly different account of the relationships between what subjects report and what they experience. This makes use of a distinction between two alternative ways that (visual) experience could unfold over time. An 'Orwellian' possibility is that a (visual) experience of an event matches what happens in the event, when it actually happens. However, this experience might be quickly forgotten and 'replaced' with a memory of having seen something slightly different. Alternatively, a 'Stalinesque' possibility is that a (visual) experience of an event lags behind the event itself, where, as above, you may end up seeing the event in a slightly different way (see Dennett, 1993 for distinction). One reason for seeing the event slightly differently is if information around the event is integrated over time; in particular, this could potentially lead to the subject experiencing a (visual) illusion.

The standard interpretation of the Sperling paradigm is an Orwellian one: subjects experience the contents of the display when it is present, but details of the display are rapidly forgotten, and subjects can only report the products of later processing (the

three letter identities that make it into working memory).

Phillips' argument is that there is also a plausible Stalinesque or 'postdictive' interpretation of the paradigm. Here, subjects have no experience of specific letters of the display before the cue, and only have letter-specific experiences of a few specific letters after the cue. Phillips argues for this using examples of other postdictive phenomena where what subjects experience and report about a stimulus depends on what one shows after it. Importantly, there are examples of postdictive effects that can occur over the kind of time-spans found in the Sperling paradigm (e.g. several hundred ms), and that can occur cross-modally (so can account for both visual and audio cues used in the paradigm).

One of these examples is the sound induced visual bounce (Sekuler et al., 1997). Here, subjects are presented with a video of two balls moving towards each other across a screen, which, when they touch, can be seen as bouncing off each other, or streaming past/through each other. If a sound is played around the point where they touch, subjects are more likely to report seeing bouncing rather than streaming. The authors report that sounds played 150ms after the point of coincidence make perception of bouncing more likely, and Choi and Scholl (2006) make this up to 200ms after coincidence. Watanabe and Shimojo (2001) also found that when a sound is played at the point of coincidence, the likelihood of seeing a bounce is affected by other sounds played up to 500ms either side of it.

So, here is a clear example of a cross-modal postdictive phenomenon with parameters similar to those found in the Sperling paradigm. Subjects' reports about the stimuli

(two moving balls) are affected by a subsequent stimulus (sound) that can occur up to 500 ms later. Phillips (2011b) also outlines an account of attention in which attending to a location amplifies whatever information has already gained from that location. This makes it possible for subjects to attend to information that is no longer physically present, or present in consciousness.

These two ideas can be used to suggest an alternative Stalinesque interpretation of the Sperling paradigm: subjects only experience specific letters in the cued rows, after the cue is shown, where cues direct attention to amplify pre-conscious representations of the letters. Importantly, this is also consistent with the existence of generic phenomenology before the cue, thus enabling Phillips to 'explain away' subjects' reports of visual richness in a similar way to the account offered earlier.

The temporal sequence of experiences is then as follows: when the display is present subjects have an experience of generic phenomenology (general letter-iness in a grid formation), which is added to several hundred milliseconds later by an experience of the specific letters in the cued row that they are asked to report on. Reports of seeing 'all the letters' stem from a combination of generic phenomenology and the fact that subjects can identify most of the letters in the cued row. Instead of a persisting experience of specific letters enabling performance at the task, short term memory (informational persistence) stores letter identities that can be accessed with an attentional cue. Having provided an alternative account of the Sperling paradigm, Phillips concludes that the onus is on proponents of the Orwellian/rich/overflow account to argue why their account is superior (something he thinks it hard to do).

The mis-match problem

One problem with the sparse/non-overflow accounts above is how the contents of various bits of processing come together in phenomenology in a way that preserves how things 'seem' to us. In particular, it looks like proponents of the sparse/non-overflowing view need to give some account of what generic phenomenology is like, how it differs from 'specific' phenomenology, and how the two might contribute to conscious experience. To illustrate this problem, Block writes that before and after the cue in the Sperling paradigm, which is supposed to mark a shift between pre-cue generic phenomenology and post-cue specific phenomenology: "…subjects report no such phenomenological shift…The vast literature on this topic…contains no mention of such a thing as far as I know. I myself can testify that even looking for such a shift, one does not experience it" (Block, 2007, p. 532).

In response, it is essential first to reiterate that according to the discussion above there is no persisting visual image of the display that lasts until the cue (visible persistence is very short-lived), so there is no continuing visual image in which a shift from generic to specific phenomenology could be marked. It could be that generic phenomenology is experienced first (when the display is first presented), and specific phenomenology of individual letters, reconstructed from memory, only comes much later (with a gap in the middle). Yet even if there is not a sudden shift in the phenomenology in ongoing experience, one might think that the different types of phenomenology would 'look' different and so be noticed anyway.

Here though one can point out that all parties face general problems in linking up

information processing, memory content, and phenomenology. Proponents of the rich/overflowing view, as above, need to give an account of why subjects do not report phenomenological changes as the contents of iconic memory change and decay over time. Proponents of the sparse/non-overflowing view of conscious content need to explain why subjects do not report on differences between generic and specific phenomenology. For both parties then, there is no straightforward link between how things appear and the contents of reports, memories, and perceptual processing. In this case, neither side is at a clear advantage from the point of view of accounting for visual phenomenology.

However, this 'mis-match' problem may be based on a 'movie screen' model of consciousness which assumes that visual phenomenology is, more or less, like a movie screen: where spatial or temporal gaps are noticeable, low-grade images are noticeably different to high-grade ones (e.g. sharp vs. fuzzy), and where parts of images have to be located and coloured and determined all the time. If one assumes a movie screen model of consciousness, then the lack of a noticeable jump between generic and specific phenomenology is a problem for the sparse/non-overflow view, because on a movie screen, it would be very obvious.

However, if one rejects the movie screen model, then the mis-match problem can be explained away. And indeed, sparse/non-overflow accounts are amenable to 'illusions' of richness via generic phenomenology, mis-matches between reports and information processing, and the mind having clever ways of dealing with gappy, changing, indefinite content as and when needed. That is, there need be no noticeable shift between generic and specific phenomenology, nor between rapidly changing

content, because perceptual experience is just not of the format where this could be noticed. We can be mistaken about how much access we have to the world at any point in time via perceptual experience.

Although the idea of illusions in consciousness are often deemed problematic, illusions or mis-matches are arguably just what you end up with if you reject a movie screen model of consciousness and take note of empirical and theoretical work on perception (see early discussions of these issues in Noë, 2002). As such, the sparse/overflow accounts described above, that reject the movie-screen model and are based on well-known perceptual phenomena, seem well placed to account for the complex relations between the contents of iconic memory and phenomenology.

One might wonder however what a rich/overflow account would look like without the movie screen assumption. Yet dropping this assumption removes much of the motivation behind the view. The primary reason for thinking that the contents of iconic memory are in the contents of phenomenal consciousness is that subjects' reports suggest as much; they report seeing 'all the letters'. Yet if visual experience is not as of a movie screen, and mismatches between reports and information processing or memory are possible, then there is no strong reason to continue to take reports at face value given conflicting empirical evidence. Instead, we allow other routes to generating these kind of reports (e.g. based on generic phenomenology and expectations), and so end up at the sparse/non-overflow view.

While more work needs to be done on exactly how these mis-matches work, explanations of the Sperling paradigm featuring generic phenomenology may be on

the winning side. However, more recently another kind of short term memory store has been proposed to contribute rich phenomenal content to consciousness, that cannot be explained away using the strategies here. This leads into deeper questions still about the nature of working memory and its relationship to phenomenal content, discussed below.

Fragile VSTM

More recently variations of change detection paradigms have been used to suggest that there is another type of memory that potentially provides a rich, and longer lasting, set of conscious content that overflows working memory (Block, 2011; also see Lamme, 2006). Fragile visual short term memory (fragile VSTM) is argued to sit between iconic memory and working memory in terms of capacity and duration (e.g. Sligte et al., 2010, 2008; Vandenbroucke et al., 2011). Because of its longer duration, post-dictive explanations can't be applied (they only work for short time spans), and while appeals to generic phenomenology can still be used, some mechanism other than informational persistence needs to be used to explain how subjects are able to identify apparently large number of letters long after they have disappeared from view.

First, it is important to understand the source of the evidence in favour of the existence of fragile VSTM. The paradigm used to identify fragile VSTM uses three conditions as follows (see also Figure 2):

Iconic-cue condition: A 'memory array' of items is displayed for around 250ms, and subjects are asked to remember as many of them as possible. A cue to a location of a particular item in the array is displayed to subjects within the temporal window used in the Sperling paradigm. After this, a probe or test display of similar objects to the memory array is shown, and subjects have to state whether the cued item has changed. As the name suggests, this condition tests the capacity of iconic memory, which is fairly high.

Retro-cue condition. As above, a 'memory array' of items is displayed for around 250ms, and subjects are asked to remember as many of them as possible. This time a cue is shown long after iconic memory has decayed (1000ms). After this, a test display is shown of similar objects to the memory array, and subjects have to state whether the cued item has changed.

Post-cue condition: As above, a 'memory array' of items is displayed for around 250ms, and subjects are asked to remember as many of them as possible. After a delay interval of up to 1000ms, a test array is displayed, and then disappears, and then a cue to a particular item location is displayed. Subjects have to state whether the cued item has changed. The post-cue condition is supposed to measure the capacity of working memory, as the presentation of the test array *before* the cue provides a lot of visual interference, so should overwrite any information that is not already in working memory.
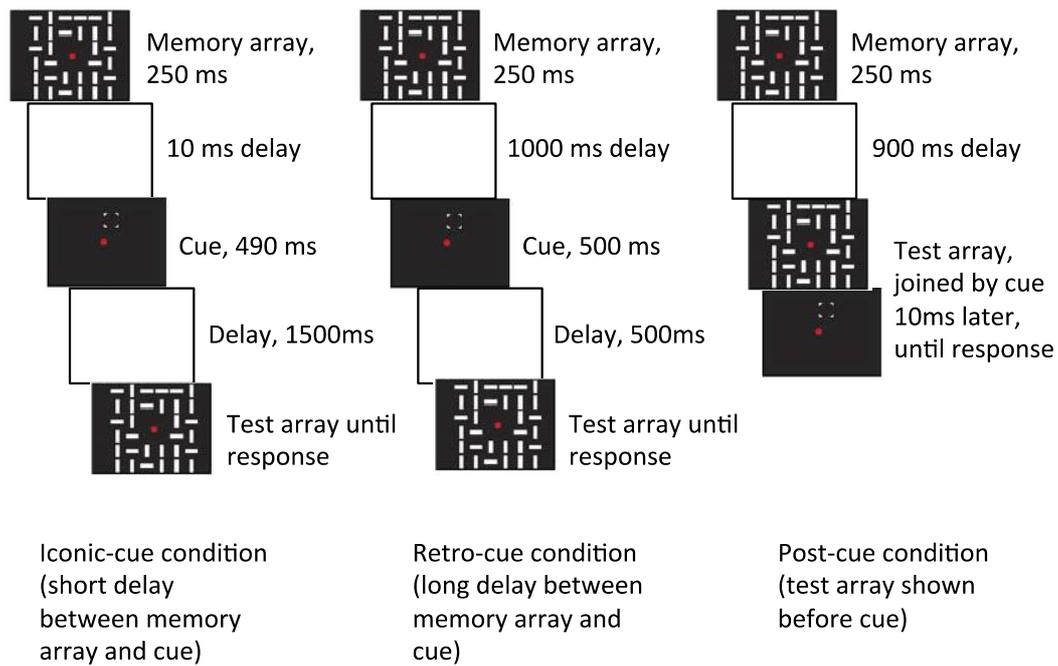
Figure 2: The basic differences between iconic-cue, retro-cue, and post-cue

conditions. Adapted from Figure 1, Sligte et al. (2008).

The interesting condition here is the retro-cue condition. The memory capacity

measured in this condition varies widely over the specifics of the task (e.g. up to 10

low-resolution simple items, down to 3 high resolution complex items, see Sligte et

al. 2008, 2010), but the capacity appears to be robustly below that of iconic memory

(measured using the iconic-cue condition), and above that of working memory

(measured using the post-cue condition). This evidence is then used as evidence of a

third type of visual memory, fragile VSTM (see Sligte et al. 2008 for early and

sustained defence of this). It is argued that subjects in these studies use their visual

experiences, supported by fragile VSTM, to perform the change detection and

identification tasks. Again then, this can be used to provide support for

rich/overflowing accounts of the contents of consciousness; subjects seem to have seen more items than they can report at any one time.

Although the strategies used above against rich/overflowing content cannot be straightforwardly applied to untangle fragile VSTM, there are other ways to challenge the idea that fragile VSTM contributes rich content to visual experience. These focus on how the capacity of fragile VSTM is measured, and if it actually forms a separate memory store to working memory after all. If it does not, then fragile VSTM cannot be used as part of a rich/overflow account of conscious content. This raises further important questions about the nature of working memory, and so of its links with consciousness.

Challenges to Fragile VSTM

The first challenge is that change detection (used by Sligte et al. 2008) is a lot easier than identification (used in the Sperling paradigm). Subjects in the change detection paradigm just have to process enough information to tell if an item has changed (e.g. changed orientation, changed colour), which a much less stringent way of assessing memory capacity than requiring subjects to freely identify letters. Phillips (2011a) raises this challenge and suggests that given that change detection has such low informational demands, it might even be the case that it can be performed successfully using unconscious priming mechanisms (pp. 405-406). Similarly, it could be the case that subjects use feelings of familiarity generated from unconscious processing to make a judgment about whether the display has changed, rather than comparing specific visual experiences.

Sligte et al. (2010) picks up this criticism and also tests change detection and an identification task with complex items (harder), instead of a change detection task only on oriented rectangles (easier), to measure the capacity of fragile VSTM. However, presumably partly to preserve the structure of the experimental paradigm, the identification task is still only a forced decision task among 4 alternatives (so still far easier than in the Sperling paradigm, in which there are 26 alternatives). Yet even in these conditions measured capacity drops sharply: capacity of fragile VSTM measured according to change detection is 4.6 items, and capacity measured according to the identification task is 3.3 items. Importantly, these capacity measures are at the high end, but within, the bounds of working memory capacity.

Further questions can be raised about measures of the capacity of fragile VSTM based on the type of stimuli used. Earlier experiments from Landman et al. (2003) using the change detection paradigm were problematic as subjects could have been 'chunking' oriented rectangles into larger spatial groups. In this case capacity measures would have massively overestimated the capacity of fragile VSTM: instead of storing 32 individual rectangles, the array could have been stored as 4 larger chunks. Sligte et al. (2008) controlled for this by rotating all the non-test items in the array by 90 degrees in the test array (intended to rule out chunking strategies), and in Experiment 3, by using four orientations rather than only two. Under these conditions, again capacity measures dropped significantly to 5.5 items, which is at the high-end of working memory capacity.

Another feature of these experiments investigated by Matsukura and Hollingworth (2011) is the amount of practice that subjects need to get up to the capacity measures found in the Sligte studies. In the initial study (Sligte 2008) subjects had 3 hours of practice, and could also select to re-do trials in the experimental phase. Matsukara and Hollingworth found that without a significant amount of practice, their subjects had much lower capacity measures, and even with 80 minutes of practice they still had lower capacity measures than those found by Sligte and colleagues. Practice and training also remain features of the later studies, with Sligte et al. (2010) giving less training time, but requiring a certain performance level before subjects engaged in the main task.

The worry here is that practice need not necessarily increase memory capacity, but can make coding, storage and retrieval more efficient, making it look like capacity has increased. In this case, subjects may be working at the high end of working memory capacity, rather than using a separate, higher-capacity memory store.

The challenges above are aimed mainly at the capacity measures of fragile VSTM, to show that they are actually within standard capacity measures for visual working memory, though often at the higher end. Yet there is still the fact that the capacity measures for the retro-cue conditions (supposed to measure fragile VSTM) are usually double those for the post-cue conditions (supposed to measure working memory). So, even if the capacity measures are not actually that high, there still seems to be evidence in favour of a separate higher-capacity memory store.

However, there are reasons to question this. Matsukara and Hollingworth (2011) also

tested whether there is a capacity difference between a partial report condition (the cueing condition the Sperling paradigm) and something similar to a full report condition (condition with no cues in the Sperling paradigm). In the partial report condition a cue is shown after the main display (as is standard), but in the full report condition here, all the items in the display are cued, effectively giving a neutral cue, but also providing visual interference. Under these conditions one would expect much higher capacity measures for the partial report condition, which assesses fragile VSTM, than the full report condition, which due to the visual interference provided should only assess the contents of working memory. The authors found a small difference between the two conditions, but the capacity limits of fragile VSTM were still within the bounds of working memory capacity (the higher capacity measure for fragile VSTM = 4.7 items).

Somewhat turning the tables in these debates, Matsukara and Hollingworth suggest that the differences in capacity measures can be explained as attentional effects within working memory itself: attention can prevent decay or interference to selected items in working memory (Makovski et al., 2008). This goes against standard conceptions of attention and working memory; that attention directs content into working memory, which provides (potentially only some of) the contents of consciousness. Here though, attention also seems to modulate the content of working memory, and so the contents of consciousness, making them more or less 'available' to other consumer systems.

Importantly, this idea that attention can work even within the contents of working memory presents a strong challenge to way that fragile VSTM is identified. Makovski (2012) and Rerko et al. (2014) show that retro-cues still benefit performance at

change detection tasks, even if the cue comes *after* visual or cognitive inference, or a switch of attention, both of which should wipe out fragile VSTM. Note that the post-cue condition is used to test working memory capacity precisely because the test array provides visual inference that wipes out anything else. So here, even in conditions that are only supposed to test the capacity of working memory, it is still beneficial to give subjects an attentional cue before the test array is shown. If these retro-cue benefits occur for content that is squarely in working memory, then higher capacity measures gained during retro-cue conditions vs. post-cue conditions (or vs. no cue or neutral cue condition) cannot be used as a criterion to separate two kinds of memory; they simply illustrate the effects of attention on the contents of working memory.

Here then, not only is the capacity of fragile VSTM not particularly high, but it is not a separate memory store from working memory. What is termed fragile VSTM is the product of efficient encoding, storage and retrieval within working memory, with storage and retrieval made yet more targeted by using an attentional cue. There is therefore no need to posit fragile VSTM as an additional memory store that 'overflows' the contents of working memory. If fragile VSTM simply doesn't exist, then its contents obviously cannot provide support in favour of the rich/overflowing view of conscious content.

What's in Working Memory?

More pressing for discussions about the relationship between short-term memory and the contents of visual phenomenology, is that there now appear to be serious questions about how to assess the capacity and contents of working memory. In

addition to the experimental work above on showing the effects of attention within working memory, it is fairly well known that measured working memory capacity is sensitive to the complexity of the items it stores (Alvarez and Cavanagh, 2004), and to the type of testing used (Makovski et al., 2010). While it is still controversial whether working memory has a limited numbers of slots or is a more divisible resource (Bays and Husain, 2008 and challenges), there seem to be several sources of variation in capacity.

This is problematic for rich/overflow views as stating that conscious content overflows working memory relies on there being a fairly foolproof way of assessing what the contents of working memory are. Yet if the capacity of working memory can be manipulated by attention, by practice, by the task and the visual stimuli used, then it is potentially quite difficult to identify what would count as convincing empirical evidence of content that overflows working memory.

It is also problematic for sparse/non-overflow views, but in a slightly different way. If working memory is defined functionally (roughly, as that which makes content available to other consumer systems, including allowing subjects to act and report on content), then whether or not we have a stable way of measuring its content is perhaps not pressing. The claim is just that the contents of working memory are the contents of consciousness, whatever the contents of working memory end up being. Yet the very same problems that affect the measurement of working memory also affect how to define its function. If the functional definition of working memory is abstract enough that these problems of measurement can be ignored, then it may become an

ambiguous and/or empty definition, and so leave theories of consciousness that involve working memory also problematically ambiguous.

Conclusion

This chapter has analysed arguments that short term visual memory provides rich conscious content that overflows the contents of working memory. These arguments were based on Sperling's partial report paradigm (iconic memory) and Sligte's change blindness studies (fragile VSTM) that appear to show that more content is stored than can be reported at any one time, and that this content is present in consciousness. It was suggested above that there are alternative explanations for subjects' behaviours and reports of richness in the Sperling paradigm that fit with what is known about the structure of iconic memory and perception more generally. In addition, it has been argued that fragile VSTM likely does not exist as a separate store to working memory. Both iconic memory and fragile VSTM are therefore poor candidates for supporting the idea that the contents of experience overflow the contents of working memory. These debates have highlighted deeper issues about how to relate the contents of information processing and memory to the contents of consciousness, and how to identify the contents of working memory in the first place; questions that both sides of the debate need to have an answer to.

And one final point that is often skated over: even if we were to accept that the contents of iconic memory and fragile VSTM are present phenomenally, it arguably would not provide much support for a truly 'rich' view of experience. If the contents of iconic memory and fragile VSTM overflow working memory, they do not do so by

very much. In the Sperling paradigm subjects can store information about 9 (out of 12) letters in iconic memory, and in the more conservative tests of fragile VSTM, capacity is around 4-5 items (less for complex objects). While having a detailed visual experience of nine letters is more impressive than experiencing only three, it does not support a picture of particularly rich visual detail. In this case, even if the arguments above fail to convince, the experimental work still presents a rather sparse picture of the contents of consciousness.

Related topics

See also 'Taxonomy and Unity of Memory', 'Phenomenology of Remembering', 'Memory and Consciousness'.

References

Alvarez, G.A., Cavanagh, P., 2004. The capacity of visual short-term memory is set both by visual information load and by number of objects. Psychol. Sci. 15, 106–111.

Bays, P.M., Husain, M., 2008. Dynamic shifts of limited working memory resources in human vision. Science 321, 851–854.

Block, N., 2011. Perceptual consciousness overflows cognitive access. Trends Cogn. Sci. 15, 567–575.

Block, N., 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. Behav. Brain Sci. 30, 481–499.

Brockmole, J., Wang, R.F., 2003. Integrating visual images and visual percepts across time and space. Vis. Cogn. 10, 853–873.

Castelhano, M.S., Henderson, J.M., 2008. The influence of color on the perception of scene gist. J. Exp. Psychol. Hum. Percept. Perform. 34, 660.

Choi, H., Scholl, B.J., 2006. Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. Percept.-Lond.- 35, 385.

Cohen, M.A., Dennett, D.C., 2011. Consciousness cannot be separated from function. Trends Cogn. Sci. 15, 358–364.

Dehaene, S., Changeux, J.-P., Naccache, L., 2011. The global neuronal workspace model of conscious access: from neuronal architectures to clinical applications, in: Characterizing Consciousness: From Cognition to the Clinic? Springer, pp. 55–84.

Dennett, D.C., 1993. Consciousness explained. Penguin UK.

Di Lollo, V., 1980. Temporal integration in visual memory. J. Exp. Psychol. Gen. 109, 75.

Dretske, F., 2007. What change blindness teaches about consciousness. Philos. Perspect. 21, 215–220.

Dretske, F., 2006. Perception without awareness.

Dretske, F., 2004. Change blindness. Philos. Stud. 120, 1–18.

Fodor, J.A., 2008. LOT 2: The Language of Thought Revisited: The Language of Thought Revisited. OUP Oxford.

Irvine, E., 2011. Rich experience and sensory memory. Philos. Psychol. 24, 159–176.

Irwin, D.E., Yeomans, J.M., 1986. Sensory registration and informational persistence. J. Exp. Psychol. Hum. Percept. Perform. 12, 343.

Kouider, S., De Gardelle, V., Sackur, J., Dupoux, E., 2010. How rich is consciousness? The partial awareness hypothesis. Trends Cogn. Sci. 14, 301–307.

Lamme, V.A., 2006. Towards a true neural stance on consciousness. Trends Cogn. Sci. 10, 494–501.

Landman, R., Spekreijse, H., Lamme, V.A., 2003. Large capacity storage of integrated objects before change blindness. Vision Res. 43, 149–164.

Luck, S.J., Hollingworth, A., 2008. Visual memory. Oxford University Press.

Makovski, T., 2012. Are multiple visual short-term memory storages necessary to explain the retro-cue effect? Psychon. Bull. Rev. 19, 470–476.

Makovski, T., Sussman, R., Jiang, Y.V., 2008. Orienting attention in visual working memory reduces interference from memory probes. J. Exp. Psychol. Learn. Mem. Cogn. 34, 369.

Makovski, T., Watson, L.M., Koutstaal, W., Jiang, Y.V., 2010. Method matters: systematic effects of testing procedure on visual working memory sensitivity. J. Exp. Psychol. Learn. Mem. Cogn. 36, 1466.

Matsukura, M., Hollingworth, A., 2011. Does visual short-term memory have a high-capacity stage? Psychon. Bull. Rev. 18, 1098–1104.

Noë, A. (Ed.), 2002. Is the visual world a grand illusion?, Journal of Consciousness Studies: Controversies in the Science & the Humanities. Imprint Academic.

Phillips, I.B., 2011a. Perception and iconic memory: what Sperling doesn't show. Mind Lang. 26, 381–411.

Phillips, I.B., 2011b. Attention and iconic memory. Atten. Philos. Psychol. Essays 204–227.

Rerko, L., Souza, A.S., Oberauer, K., 2014. Retro-cue benefits in working memory without sustained focal attention. Mem. Cognit. 42, 712–728.

Sekuler, R., Sekuler, A.B., Lau, R., 1997. Sound alters visual motion perception. Nature 385, 308.

Sligte, I.G., Scholte, H.S., Lamme, V.A., 2008. Are there multiple visual short-term memory stores. PLOS One 3, e1699.

Sligte, I.G., Vandenbroucke, A.R., Scholte, H.S., Lamme, V., 2010. Detailed sensory memory, sloppy working memory. Front. Psychol. 1, 175.

Sperling, G., 1960. The information available in brief visual presentations. Psychol. Monogr. Gen. Appl. 74, 1.

Tye, M., 2009. Consciousness revisited: Materialism without phenomenal concepts. Mit Press.

Tye, M., 2006. Nonconceptual content, richness, and fineness of grain, in: Gendler, T.S., Hawthorne, J. (Eds.), Perceptual Experience. Oxford University Press, Oxford.

Vandenbroucke, A.R., Sligte, I.G., Lamme, V.A., 2011. Manipulations of attention dissociate fragile visual short-term memory from visual working memory. Neuropsychologia 49, 1559–1568.

Watanabe, K., Shimojo, S., 2001. When sound affects vision: effects of auditory grouping on visual motion perception. Psychol. Sci. 12, 109–116.