Afterimages: Their Neural Substrates and Their Role as Short-Term Memory in the Neural Computation of Human Vision

Anonymous Author(s)
Affiliation
Address
email

Abstract

Afterimages are seemingly simple yet very intriguing visual phenomena. Presently, essentially all the textbooks in vision science and in perceptual psychology introduce these phenomena; meanwhile they also ubiquitously subscribe to an incorrect view that afterimages are due to some peripheral adaptation mechanisms occurring in the retina of the eye. The contrasting view is that afterimages originate in the brain: This view is not new at all, but only recently there has been accumulating a multitude of evidence pointing to its truthfulness. Two recent and critical lines of advances related to afterimages in vision science are as follows: 1. LeVay et al. (1985) discovered a representation of the physiological blind spot in Layer 4 of the cortical area V1 (hereafter, V1-L4) in the macaque monkey's brain, and Adams et al. (2007) discovered the same in the human brain; 2. Wu (2024) re-discovered the phenomenon of an observer seeing their own blind spot as an afterimage and correlated this phenomenon to the above neuroanatomical findings. Together, these advances essentially pinpoint the first-stage neural substrate for afterimages to V1-L4. Here we build upon these advances and establish a neural theory of afterimages consisting of the following tenets: 1. Positive and negative afterimages share the same neural substrate; 2. Afterimages should be viewed as short-term memory (STM) in the brain instead of as peripheral adaptation in the retina; 3. In terms of the neural computational architecture of any cortical area, STM is sandwiched between a feedforward neural network and a feedback counterpart—it may play a computational role for variable binding. Finally, we discuss potentially fruitful bi-directional interactions between perceptual & neuroscientific researches in biological vision on the one hand and computational & engineering endeavors on artificial vision on the other.

1 Introduction

Under certain visual conditions, when a viewer sees a stimulus, they may continue to see an image of the stimulus even after the physical stimulus has already disappeared: This visual phenomenon is known as an afterimage. A basic and prominent issue pertinent to afterimages is where they occur in the human visual system: Are they merely some adaptation mechanisms happening in the retina of the eye (hereafter, this view will be referred to as the Retinal View)? or are they a form of visual memory residing in the brain (hereafter, the Brain View)? Table 1 lists some major publications concerning afterimages since the time when Newton (1691) communicated his observation of afterimages: As we can see, there have been proponents of both of these two views. In the present paper, we will demonstrate that the Retinal View is erroneous and only the Brain View is correct.

Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

Table 1: Major publications on the Retinal View vs. the Brain View with regard to afterimages

Investigators	Phenomena / Arguments	The Retinal View	The Brain View
Newton (1691)	He observed interocular transfer of afterimages and briefly suggested: "It [afterimage] seems rather to consist in a disposition of the sensorium [the part of the brain for sensation] to move the imagination strongly" (p. 154).		✓
Darwin (1786)	Adaptation: analogy between the retina and the muscular system	✓	
Binet (1886)	Interocular transfer (pp. 43-45)		✓
Delabarre (1889)	Interocular differences of seeing the same afterimage	✓	
Craik (1940)	Retinal anoxia ("blinding" an eye by finger-pressing it) disrupts afterimage.	✓	
Weiskrantz (1950)	Afterimage may be produced from imagination.		✓
Urist (1958)	Pushing an eye's ball, the eye's scene view shifts in position but its afterimage stays.		✓
Loomis (1972)	Afterimages from long-duration light stimulation with little bleaching are correlated with visual appearance.		✓
LeVay et al. (1985)	Discovery of a cortical representation of the physiological blind spot in V1-L4 in the macaque monkey's brain		
Shimojo et al. (2001)	Filling-in visual surface may generate afterimage.		✓
Tsuchiya and Koch (2005)	Interocular influence on afterimage formation		✓
Adams et al. (2007)	Discovery of a cortical representation of the physiological blind spot in V1-L4 in the human brain, providing the neuroanatomical basis for Wu (2024)		
Shevell et al. (2008)	Interocular misbinding of color and form		✓
Zaidi et al. (2012)	Physiological recordings in the macaque monkey's brain	✓	✓
Dong et al. (2017)	The Breese effect (Breese, 1899): binocular rivalry between two eyes' afterimages slower than that with physical stimuli.	✓	✓
Kronemer et al. (2024)	Shared mechanisms between afterimages and visual imagery		✓
Wu (2024)	Re-discovery of the La Hire phenomenon (i.e., seeing the physiological blind spots as afterimages), pinpointing such afterimages to the neural substrate V1-L4.		✓
Kittikiatkumjorn et al. (2025)	Afterimage color is factored by color constancy		✓

The topic of afterimages is universally taught in all the textbooks in vision science (e.g., Palmer, 1999, pp. 105 & 109) and in perceptual psychology (such a course may be known as "Sensation and Perception"; e.g., Wolfe et al., 2021, pp. 153-155). About a decade ago, essentially all such textbooks had subscribed only to the Retinal View. Presently, the situation is changing—for example, citing Zaidi et al. (2012), Wolfe et al. (2021) advocate a hybrid view as follows: "Adaptation occurs at multiple sites in the nervous system, though the primary generators are in the retina" (p. 155): The Retinal View is still a component in this hybrid view; hopefully, the Retinal View would become totally abandoned in another decade or so.

As shown in Table 1, the Brain View regarding afterimages' localization in the human visual system is not new at all: Newton (1691) was already suggesting it. However, only in the last several decades, there have been accumulating many lines of evidence in support of the Brain View. In this respect, two particularly relevant and critical findings are as follows: (1) LeVay et al. (1985) delineated a representation of the physiological blind spot in Layer 4 of the primary visual cortex (also known as the cortical area V1; hereafter, Layer 4 of V1 will be referred to as V1-L4) in the macaque monkey's brain, and Adams et al. (2007) found the same in the human brain; (2) Wu (2024) re-discovered the phenomenon of a human observer being able to see their own physiological blind spot as an afterimage and correlated this phenomenon to the neuroanatomical finding by VeLay et al. (1985) and Adams et al. (2007). Together, these recent advances decisively and precisely pinpoint the first-stage neural substrate of afterimages to V1-L4.

In this paper, we will build upon the above-mentioned recent advances and establish a neural theory of afterimages consisting of the following tenets: 1. Positive and negative afterimages share the same neural substrate: The first-stage is V1-L4, and the subsequent stages are the layer 4s in other visual cortical areas—in this respect, we will substantiate the Brain View about afterimages into a concrete form; 2. Afterimages constitute a form of short-term memory (STM) in the brain; 3. In terms of the neural computational architecture of the brain, for each cortical area, STM is sandwiched between a feedforward neural network and a feedback counterpart—it may play a computational role for variable binding. Finally, we discuss potentially fruitful bidirectional interactions between perceptual & neuroscientific researches in biological vision on the one hand and computer science & engineering endeavors in artificial / machine vision on the other.

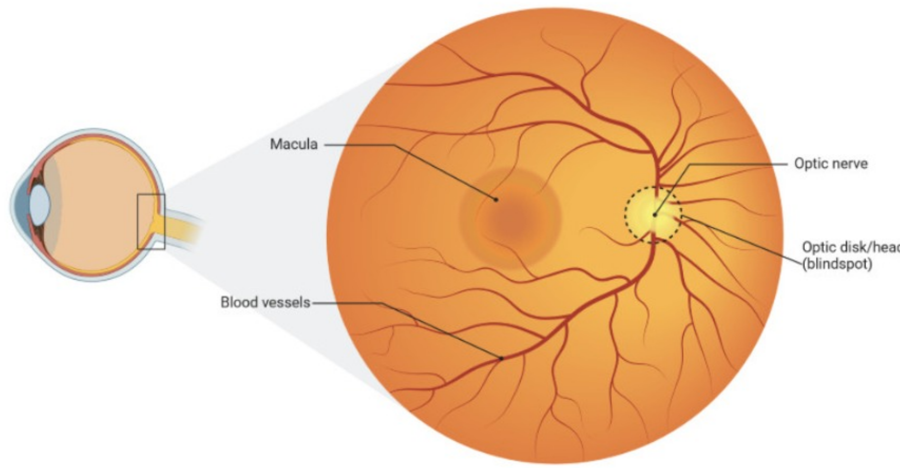
2 Physiological blind spots, afterimages, and neural localization

In this section, we will expand on the conference presentation by Wu (2024) at the Annual Meeting of the Psychonomic Society last year. As a systematic explanation regarding the relation between physiological blind spots and afterimages, this section is necessary for a complete understanding of why it is possible to deterministically pinpoint the first-stage neural locus of afterimages to V1-L4. (We understand that Wu (2024) is only an abstract and that he has not yet published his conference presentation as a full paper, but here we do acknowledge that the basic idea in this section comes from him.)

2.1 The physiological blind spot in the eye

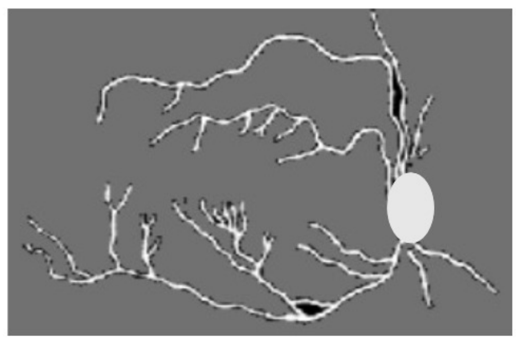
We have a physiological blind spot in each of our eyes: It corresponds to a port of the eye's retina (anatomically known as the "optic disk") where no photoreceptors (i.e., rods and cones) exist, where optic nerve fibers exit the eye, and where blood vessels enter and exit the eye (i.e., arteries entering and veins exiting the eye). This anatomical feature of the eye is clearly seen in Figure 1(a) which shows an image of a human eye's retina as seen by an ophthalmologist (i.e., eye doctor) when examining someone else's eye with some retina imaging device. The shortened term "blind spot" may mean various things in different contexts; hereafter, we will use it to refer specifically to the physiological blind spot in the eye.

The blind spot was discovered by the French scholar Edme Mariotte in the 1660s: It is certainly an amazing scientific discovery. Mariotte's method demonstrating the blind spot, however, is about how to map it within the viewer's visual field, not about how to (consciously) see it. Presently, all the textbooks in perceptual psychology, vision science, neuroscience, and ophthalmology, when mentioning about the blind spot, describe this method only (e.g., Wolfe et al., 2021, p. 40).



(a) The retina of a human eye as seen by an ophthalmologist with an eye / retina imaging device. Optic Disk = Blind Spot (BS); the image of retinal blood vessels: Purkinje Tree (PT)

(b) The subject seeing their own eye's BS & PT in positive form

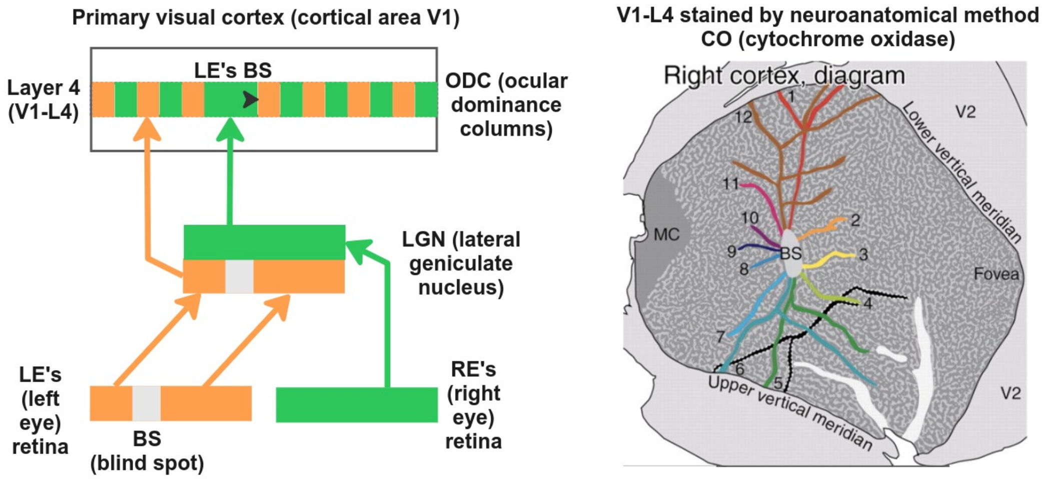


(c) The subject seeing their own eye's BS & PT in negative form

Figure 1: The blind spot and blood vessels within an eye as seen by ophthalmologist from outside and by the subject (the owner of the eye) within entopic vision.

Under special conditions, it is actually possible for the subject (i.e., the owner of the eye; we may also refer to him/her as the viewer or observer) to see their own blind spot in each eye, literally seeing the blind spot as a black hole on a lighter background or a white hole on a darker background, as illustrated in Figure 1(b) and © respectively—more generally, a colored spot on a background of the spot's complementary color; the BS may, or may not, be accompanied by the Purkinje Tree (PT) which denotes the image of retinal blood vessels. As far as we have been able to trace back, this phenomenon was first reported by the French scholar Philippe de La Hire (1640-1718) in La Hire (1694): Henceforth, we will refer to this phenomenon as the La Hire phenomenon. It was subsequently re-discovered by the Czech scientist Johann Evangelist Purkinje (1787–1869) in Purkinje (1819): Figure 1(b) is his drawing of his observation of his right eye's blind spot and retinal blood vessels. More broadly, Purkinje referred to a set of visual phenomena of a viewer seeing some characteristics of the human visual system's internal organization as "subjective vision"—presently, they are known as "entoptic vision"; therefore, the La Hire phenomenon is an instance of entoptic vision.

As mentioned by Helson (1929, pp. 352–353) and Brøns (1939, Chapter IV), many German psychologists had investigated the La Hire phenomenon before World War II. After the war, it appears that the vision research community has largely forgotten about this interesting visual phenomenon.



(a) Three early stages in the human visual system: the retinae in the eyes, LGN in the thalamus, and V1

(b) Representations of the blind spot and retinal blood vessels in V1-L4

Figure 2: Three early stages of the human visual system: retina, LGN, and V1-L4.

2.2 Seeing the blind spot as a negative afterimage

Wu (2024) described his observation of seeing an eye's blind spot as a negative afterimage—that is, normally, the blind spot would appear as a black spot on a neutral background; but under special conditions, as illustrated in Figure 1(c), it may appear negatively as a brighter spot on a darker background.

Actually, both Helson (1929) and Brøns (1939) had mentioned the possibility of seeing the blind spot as a negative afterimage: It is one of the characteristics of the La Hire phenomenon—in this respect, on the phenomenological side, Wu (2024) is a re-discovery; nonetheless, he correlated this phenomenological feature of the La Hire phenomenon and determined the neural locus of afterimages: We will describe this correlational reasoning below.

2.3 Localizing the neural substrate of the blind spot and its afterimage

Figure 2(a) illustrates the early stages of the human (or more broadly, the primate) visual system consisting of the retina, the lateral geniculate nucleus (LGN) of the thalamus, and the primary visual cortex (i.e., V1). The cortical sheet comprises six layers, with Layer 4 receiving thalamic inputs (i.e., optic radiations in the case of V1). Hereafter, we will denote this layer as V1-L4. Please note that "Layer 4" in V1 has been incorrectly labeled as "Layer 4C" in many textbooks (e.g., Wolfe et al., 2021, p. 71); see Boyd et al. (2000) and Balaram et al. (2014) for the relevant neuroanatomical evidence as to why it should be labeled as "Layer 4" instead of "Layer 4C".

The La Hire phenomenon is a wonderful psycho-anatomical means: As a matter of fact, several neuroanatomical studies have precisely localized a representation of the blind spot in V1-L4: LeVay et al. (1985) and Adams et al. (2007) (from Prof. Jonathan Horton's lab at University of California, San Francisco) are the two milestone discoveries in this regard, with the first one in the macaque monkey's brain and the second in the human brain. Though in two species and using different chemical staining methods, their central findings are essentially the same: There is a representation of the blind spot in V1-L4. For better illustration, Figure 2(b) shows a diagram of V1-L4 from a monkey's brain studied by Prof. Horton's lab. Please note that V1-L4 is a "bi-monocular" structure in the sense that for each and every tiny patch of the viewer's binocular visual field, the monocular image (i.e., ocular dominance column or ODC) from one eye resides, side by side, with that for the other eye. In Figure 2(b), white stripes and areas depict neural tissue regions in V1-L4 predominantly connected with the eye containing the blind spot, whereas the black stripes and areas depict that connected with the other. From this diagram, we should understand that the representation of the blind spot in V1-L4 does not create any physical "hole" in this neural tissue—instead, the area is invaded and occupied by the input from the other eye.

Beyond V1-L4, is there any other neural structure(s) in the primate visual system that may contain representations of the blind spot? David Hubel and Torsten Wiesel's pioneering exploration of the feline and the primate visual brains had long established that neurons in V1-L4 are primarily monocular whereas that beyond V1-L4 are mainly binocular (e.g., Hubel & Wiesel, 1968). As we already stated, each eye's blind spot is specific to that eye (i.e., monocular); therefore, the answer to this question is negative. Correlating the La Hire phenomenon with such neuroanatomical studies, we can conclude that visual sensation is represented in V1-L4. Please note that without knowing the La Hire phenomenon, we cannot argue that the blind spot representations seen in V1-L4, and this layer in general, are directly correlated with visual sensations and afterimages—in other words, one may argue that such representations are just for sub-consciousness neural activation. With the knowledge of the La Hire phenomenon, then, we can indeed pinpoint the neural substrate for visual sensations and afterimages to V1-L4. Please note that the authors of the relevant neuroanatomical studies did not link their findings with any visual phenomenon; the correlation between the afterimage phenomenon and neuroanatomical underpinnings was advanced by Wu (2024).

3 Positive and negative afterimages

3.1 The Franklin effect

One conceptual blocker to thinking of afterimages as a form of memory is that an afterimage can manifest itself as a positive or negative one. Presently, the prevailing conception about positive and negative afterimages is that they are due to different physiological causes. For instance, De Valois and De Valois (1997) suggested that positive afterimages is due to temporary persistence of discharges of ganglion neurons in the retina, whereas negative afterimages is partly due to retinal photopigment bleaching and partly due to neural adaptation at an opponent-colors stage. Likewise, Gregory (2004, p.15) presents essentially the same conception about positive and negative afterimages. The conception that positive and negative afterimages are due to separate physiological processes (or mechanisms) is schematically illustrated in Figure 3(a).

This above conception, however, is just a misconception—this is because positive and negative afterimages can be mutually converted into one another; more specifically, an afterimage can appear either positive or negative depending on whether the observer is viewing it with his/her eyes closed or open; and with the eyes open, depending on whether projecting the afterimage onto a dark, gray, or white background. This phenomenon was first observed and described by Benjamin Franklin (1706—1790) [whose life was simply too many-splendored: a founding father of the United States, a successful businessman, a scientist famous for taking electricity from the sky, and an inventor]. On June 2, 1765, in a letter to Lord Kames, Franklin described his following observation: "A remarkable circumstance attending this experiment, is, that the impression of forms is better retained than that of colors; for after the eyes are shut, when you first discern the image of the window, the panes appear dark, and the cross bars of the sashes, with the window frames and walls, appear white or bright; but, if you still add to the darkness in the eyes by covering them with your hand, the reverse instantly takes place, the panes appear luminous and the cross bars dark. And by removing the hand they are again reversed." (Franklin, 1765, p.380). This phenomenon had been further studied by Robert Waring Darwin (1766—1848), the father of the evolutionist Charles Darwin (1809—1882). On March 23, 1786, Darwin read a paper on "ocular spectra" (that was the term for afterimages at that time) before the Royal Society of London; the paper was subsequently published as Darwin (1786).

The above phenomenon has been named the Franklin effect (Roeckelein, 2006, p. 649)—but to a large extent, somehow unfortunately, it remains largely unknown besides being mentioned in some comprehensively and meticulously compiled works, such as those by Roeckelein. As illustrated in Figure 3(a), the hypothesis that positive and negative afterimages are due to different physiological processes cannot account for the Franklin effect; therefore, we would need to seek other explanations for positive and negative afterimages.

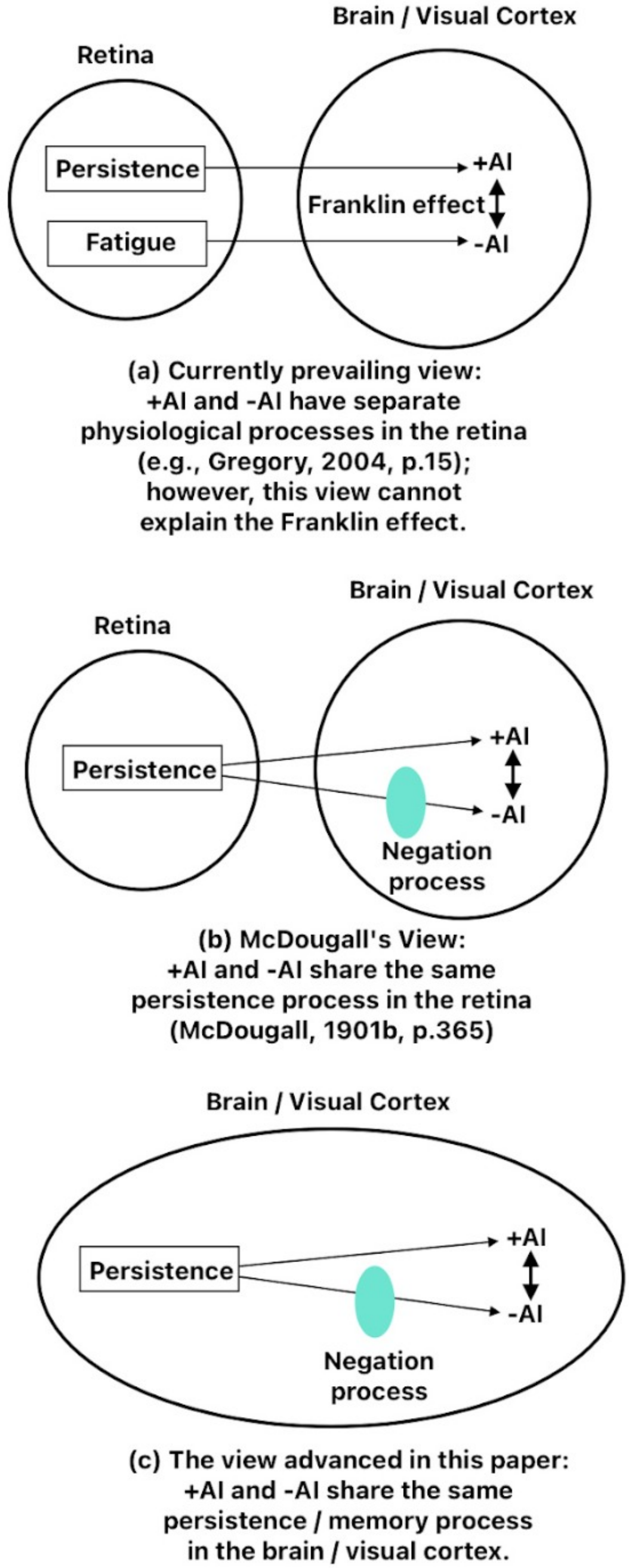


Figure 3: Three views regarding positive and negative afterimages

3.2 McDougall's view regarding positive and negative afterimages

One and a quarter centuries after Franklin and Darwin, McDougall (1901a, 1901b) rediscovered the Frank effect, and then he further claimed: "...all afterimages, negative and positive, same-colored and complementary-colored alike, are primarily due to the persistence in the retina of X-substances ..." (McDougall, 1901b, p.365). As illustrated in Figure 3(b), his claim consists of two parts: (1) positive and negative afterimages are both due to some material persistence in our visual system; (2) this persistence resides in the retina of the eye. In the previous section, however, we have already dismissed afterimage's retinal origin; therefore, we can now adopt McDougall's view about positive and negative afterimages and modify it to become the view illustrated in Figure 3(c).

4 Afterimages as visual short-term memory (STM)

Now we have established two facts: afterimage is cortical in origin, and positive and negative afterimages are both due to neural persistence. In our opinion, a neural persistence process (or mechanism) in the brain should better be conceived as visual STM—as a matter of fact, an elementary form of this view had already been suggested by Newton in 1704.

During the years 1661—1664, when Newton was an undergraduate student at Trinity College, he kept a notebook which has passed down in history and is currently in archive at University of Cambridge Library (see McGuire & Tamny, 1983). The notebook contains a section under the heading "some philosophical questions"—there, Newton wrote down a wide range of observations and questions in natural philosophy, some of them belonging to perceptual and cognitive psychology as we know today, within the topics ranging from vision, audition, memory, imagery, to consciousness. For vision, he recorded a number of visual phenomena: One of them is a particular form of positive afterimages, as he described in this way: "There is required some permanency in the object to perfect vision, thus a coale whirled round is not like a coale but fiery circle . . . " (McGuire & Tamny, 1983, p. 387). About 40 years later, when Newton (1704) published his book "Opticks", he did compile this visual phenomenon as one of the queries appended near the end of the book: "Query 16: . . . And when a Coal of Fire moved nimbly in the circumference of a Circle, make the whole circumference appear like a Circle of Fire; is it not because the Motion excited in the bottom of the Eye by the Rays of Light are of a lasting nature, and continue till the Coal of Fire in going round returns to its former place?" (see McGuire & Tamny, 1983, p. 237)

What Newton was describing as "permanency" in his Trinity College notebook and as "lasting nature" in his book "Optiks" happens within the observer's mind—apparently, it is a form of memory in the human brain. Furthermore, Newton did point out that this form of positive afterimages plays an active functional role in color perception—specifically, in temporal color summation (also known as color mixture or color fusion)—see "Persistence of Vision" (2025).

Now, once we understand that afterimages play an active computational role in human vision, should we conceptualize afterimages better as STM than merely as adaptation? We suggest and believe that the answer is "yes".

5 Interdisciplinary Interactions between Cognitive Science & Neuroscience and Computer Science & Engineering

If one learns about afterimages from an introductory textbook in vision science, one would easily get an impression that all about afterimages have already been researched and known. This is very far from the truth. In the present paper, we have witnessed that on the one hand there have been a vast array of observational and experimental data about afterimages accumulated over the last 350 years or so (after Newton's recording of his observations on afterimages around 1664), and on the other, some basic questions concerning afterimages have not yet been fully answered. We have summarized some recent advances and theorized some functional or computational aspects about afterimages, but our paper certainly has limitations: One of them is that we have not yet established a quantitative model encompassing all or most of the empirical data—in this regard, we believe that interactions between cognitive science & neuroscience on the one hand and computer science & engineering on the other would become fruitful—such interactions may eventually to solve some fundamental problems in neuroscience, including those basic questions surrounding afterimages.

As narrated by Sejnowski (2018), Geoffrey Hinton, the godfather of deep learning, was closely associated with cognitive science neuroscience at UCSD, CMU, and UCL before his distinguished tenure at Google. Sejnowski himself was also in the same kind of interdisciplinary research environment and was closely connected with Hinton (Sejnowski, 2018, p. 60); as a pioneer and prominent computational neuroscientist, Sejnowski's research interests and research certainly lean more on the neuroscience side: As he summons in his book, there is still a vast scientific landground about human perception, cognition, and intelligence yet to be explored.

6 Conclusions

Currently, the standard textbook teaching about afterimages is that they origin in the retina of the eye. Here we have presented an array of evidence—particularly the phenomenon of the blind spot visible as an afterimage—to argue for cortical origins of afterimages. Furthermore, we have developed a new theoretical perspective for understanding afterimages in human vision, consisting of the following tenets: 1. Positive and negative afterimages share the same neural substrate; 2. Afterimages should be viewed as short-term memory (STM) in the brain instead of as peripheral adaptation; 3. In terms of the neural computational architecture of the brain, this STM is sandwiched between a feedforward neural network and a feedback counterpart—it may play a role for variable binding.

NeurIPS is a great venue for interdisciplinary interactions between cognitive science & neuroscience on the one hand, and computer science & engineering on the other. Our paper belongs to the former—hopefully, it would attract some talents and efforts from the latter to tackle the scientific issues raised here in one direction and to transfer and incorporate some of our ideas proposed here into real-world applications in the other.

References

Adams, D. L., Sincich, L. C., & Horton, J. C. (2007). Complete pattern of ocular dominance columns in human primary visual cortex. Journal of Neuroscience, 27, 10391–10403.

Binet, A. (1886). La psychologie du raisonnement: Recherches expérimentales par l'hypnotisme. Publisher: Félix Alcan.

Breese B. B. (1899). On inhibition. Psychological Monographs, 3, 1–65. https://doi.org/10.1037/h0092990

Craik, K. J. W. (1940). Origin of visual after-images. Nature, 145, 512.

Darwin, R. W. (1786). New experiments on the ocular spectra of light and colours. Philosophical Transactions of the Royal Society of London, 76, 313-348. https://doi.org/10.1098/rstl.1786.0016

Delabarre, E. B. (1889). On the seat of optical afterimages. The American Journal of Psychology, 2, 326-328. https://doi.org/10.2307/1411810

De Valois, R.L., & De Valois, K.K. (1997). Neural coding of color. In A. Byrne & D.R. Hilbert (Eds.), Readings on color—Vol. 2: The science of color (pp. 94-140). Cambridge, MA: MIT Press.

Dong, B., Holm, L., & Bao, M. (2017). Cortical mechanisms for afterimage formation: evidence from interocular grouping. Scientific Report, 7, 41101. https://doi.org/10.1038/srep41101

Franklin, B. (1765). Letter to Lord Kames. In A. H. Smyth (1970, Ed.), The writings of Benjamin Franklin: Volume IX (pp. 1783–1788). New York: Haskell House.

Gregory, R.L. (2004). Oxford companion to the mind. Oxford: Oxford University Press.

Kittikiatkumjorn, N., Phusuwan, W., Sena, A., Kasemsantitham, A., & Chunamchai, S. (2025). Color afterimage is based on the color you perceive rather than the actual color of the object. Abstracts of the 25th Annual Meeting of the Vision Science Society (VSS), 25, 190.

Kronemer, S. I., Holness, M., Morgan, A. T., Teves, J. B., Javier Gonzalez-Castillo, J., Handwerker, D. A., & Bandettini, P. A. (2024). Neuroscience of Consciousness, 2024, niae032. doi: 10.1093/nc/niae032

La Hire, P. de. (1694). Dissertation sur les différents accidents de la vue. In Mémoires de Mathématique et de Physique (pp. 233–302). Anisson.

LeVay, S., Connolly, M., Houde, J., & Van Essen, D. C. (1985). The complete pattern of ocular dominance stripes in the striate cortex and visual field of the macaque monkey. Journal of Neuroscience, 5, 486–501.

Loomis, J. M. (1972). The photopigment bleaching hypothesis of complementary after-images: a psychophysical test. Vision Research, 12, 1587–1594.

McDougall, W. (1901a). Some new observations in support of Thomas Young's theory of light and color-vision: Part I. Mind, 10, 52–97.

McDougall, W. (1901b). Some new observations in support of Thomas Young's theory of light and color-vision: Part III. Mind, 10, 347-382.

McGuire, J. E., Tamny, M. (1983). Certain philosophical questions: Newton's Trinity notebook. Cambridge University Press.

Newton, I. (1691). Letter to John Locke. In Turnbull, H. W. (1961, eds.). The Correspondence of Isaac Newton, Vol. 3, 1688–1694. Cambridge University Press.

Palmer, S. E. (1999). Vision science: photons to phenomenology. The MIT Press.

Persistence of Vision (2025). In Wikipedia. https://en.wikipedia.org/wiki/Persistence$_{o}f_{v}ision$

Purkinje, J. (1819). Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht. Prague: Vetterl.

Roeckelein, J. E. (2006). Elsevier's dictionary of psychological theories. Amsterdam: Elsevier.

Sejnowski, T. (2018). The deep learning revolution. The MIT Press.

Shevell, S. K., St Clair, R., & Hong, S. W. (2008). Misbinding of color to form in afterimages. Visual Neuroscience, 25, 355-360. DOI: 10.1017/S0952523808080085

Shimojo, S., Kamitani, Y., & Nishida, S. (2001). Afterimage of perceptually filled-in surface. Science, 293, 1677–1680.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. Nature Neuroscience, 8, 1096-1101. doi: 10.1038/nn1500

Urist, M. J. (1958). Afterimages and ocular proprioception. American Medical Association Archives of Ophthalmology, 160, 161-163.

Weiskrantz, L. (1950). An unusual case of after-imagery following fixation of an "imaginary" visual pattern. Quarterly Journal of Experimental Psychology, 2, 170-175. https://doi.org/10.1080/17470215008416594

Wolfe, J., Kluender, K., Levi, D., Bartoshuk, L., Herz, R., Klatzky, R., Merfeld, D. (2021). Sensation and perception (6th ed.). Oxford University Press.

Wu, C. Q. (2024). The phenomenon of seeing one's own blind spots as afterimages and its implication for the cortical origin of afterimages. Abstracts of the Psychonomic Society 65th Annual Meeting, 29, 232.

Zaidi, Q., Ennis, R., Cao, D., & Lee, B. (2012). Neural locus of color afterimages. Current Biology, 22, 220-224. DOI: 10.1016/j.cub.2011.12.021

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract of our paper contains a summary of the contributions and scope of this paper.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: This paper contains a section discussing open questions and the potentiality of collaboration between scientists and engineers in the related domains to tackle such questions.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes]

Justification: The each theoretical claim, we have clearly provided its assumptions and the relevant perceputal and neuroscientific evidence in support of it.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [NA] .

Justification: This paper does not contain computational experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [NA] .

Justification: This paper does not contain computational experiments.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA] .

Justification: This paper does not contain model training.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [NA] .

Justification: This paper does not contain computational experiments.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [NA] .

   Justification: This paper does not contain computational experiments.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We have carefully checked the NeurIPS Code of Ethics and declare that our research confirms to the Code.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [NA] .

    Justification: This paper does not contain potential societal impacts.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA] .

    Justification: This paper does not contain computational data or models that have a high risk for misuse.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [NA] .

    Justification: This paper does not contain new computational assets.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA] .

    Justification: This paper does not contain new computational assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA] .

    Justification: This paper does not contain new experiments that involve human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA] .

    Justification: This paper does not contain new experiments that involve human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA] .

Justification: This paper does not involve LLMs.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.