Visual Secrets: a human security primitive

Enka Blanchard, Sébastien Bouchard, Ted Selker

Visual Secrets: a human security primitive

ENKA BLANCHARD, Digitrust, Loria, Université de Lorraine, France
SÉBASTIEN BOUCHARD, LaBRI, Université de Bordeaux, France
TED SELKER, University of Maryland, Baltimore County, USA

This article presents and evaluates an idea for a non-transferable secret that can be used for security verification. This new type of security primitive relies on the pre-semantic treatment of images in the human brain. By showing users an image for a limited time, we show that they can find it again when it is shown among a larger set. Despite their ability to recognise their image, they cannot reliably communicate to someone else exactly how to do so. As the secret is embedded in the very act of recognition, it cannot be shared by the user — whether voluntarily or through coercion. We report on the initial results of a usability study on 151 subjects which showed that subjects can recognise their image shown among 20 similar images with an accuracy of 79% to 86%, compared with an expected baseline of 5%. Despite their recognisability, the 'secret' images were hard to describe in unambiguous ways: no assessor managed to accurately identify the images from the description given by the subjects.

Additional Key Words and Phrases: Usable security, Biometrics, Cognitive psychology, image recognition, Verifiability

ACM Reference Format:
Enka Blanchard, Sébastien Bouchard, and Ted Selker. 2018. Visual Secrets: a human security primitive. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/1122445.1122456

1 DEFINING VISUAL SECRETS

Researchers in usable security often talk about "something you are, know or have". Those secrets are often shareable: one can give their home keys to a friend, be coerced into revealing their passwords, or have their fingerprints stolen [15]. The problem we seek to tackle is the following: is it then possible for humans to have secrets that cannot be shared?

Let us suppose two individuals decide to meet in public and want to be able to ascertain each other's identity. However, they are afraid of one of them being coerced into revealing the identification mechanism, and being replaced by an adversary. Any passphrase or callsign could be obtained under coercion and replicated. The problem then is to find a secret that they could recognise but would not be able to share, no matter the context.

In a formalised version, this problem is not *a priori* solvable by independent agents in a classical computing setting. As Turing Machines can simulate each other, any communication between agents would be indistinguishable if one agent were simulated. This is not necessarily true in a quantum setting, as a non-simulatable protocol could potentially be found thanks to the no-cloning theorem — depending on the formalism used [22].

However, humans are not Turing Machines, and their physical bodies have unique abilities. One option is to use something that only they could do, for example a behavioural biometric. This is possible in the abstract case, but multiple problems exist with those, from high error rates to biometric identity theft [6, 17]. Moreover, this type of secret can require complex apparatus to measure.

A second lead is then to use our specialised cognitive functions. One such function of particular interest to us is linked to image recognition. As has been demonstrated since the 1960s, humans have an extensive memory for visual stimuli [13, 16, 23, 25]. This has already been used as a source of security primitives, for example with authentication in the case of visual passwords [11], as well as with biometric methods [1, 2, 26]. Most importantly in our case, a significant aspect of this image recognition happens in a pre-semantic and pre-cognitive fashion, and requires no conscious effort, thanks to specialised neural pathways in multiple areas of the brain [13, 18]. This is related to the difference between recognition and recall [10]. The mind's pre-semantic treatment means that there might be a loss of information during image recognition. As such, the ability to recognise an image is not directly related to our mental description of it, and any description might ignore some key elements of the picture.

We take inspiration from both this cognitive science research and concepts from zero-knowledge proofs [9]. We thus use this pre-semantic treatment as a source of secrets that are recognisable but not shareable, and call the resulting primitive a *visual secret*[1]. A user with unlimited time and good eyesight might be able to describe exhaustively each pixel of an image. However, practical protocols would have reasonable constraints on the time spent describing images. This paper explores the viability of this new primitive, and the relationship between describability and recognisability.

1.1 Main results

We present two main contributions in this paper:

- we introduce and formalise the concept of visual secrets;
- we report on the results of a usability study on 151 subjects that demonstrates the viability of visual secrets: they have both high recognisability and low describability.

2 EMPIRICAL STUDY

2.1 Protocol

The online protocol tested image recognition and describability for three images series. It was split into four sections:

(1) A single introductory page informing subjects of their rights (including the right to quit at any point) and informing them that they would have to confirm at the end to submit the experimental data. It also asked whether they had performed or seen someone else perform the experiment and whether they were on a mobile device.
(2) Three pages, each featuring a picture (one per series). The instructions were to describe the picture in at most 10 words to try to make it identifiable among similar images.
(3) Three pairs of pages, each page featuring 2 rows of 5 images. Each pair of pages corresponds to a series of 20 images. These images were randomly distributed between the two pages, in such a way that all images were shown exactly once in this section. Subjects were asked to select an image if they thought it was one of those they had seen earlier, and could also select "none".
(4) A conclusion page thanking them for their input, indicating their scores on the memory phase, and asking them to confirm the submission of the experimental data.

We used A/B testing to randomly assign the order of the first two image series. The third series was always shown last, and the image recognition order was the same as the presentation order. The tasks of writing the second and third images descriptions were used as distractor tasks in order to limit the effect of short-term memory. The A/B testing allowed us to measure any order effect (which was not statistically significant).

[1]There is no relationship between visual secrets and visual cryptography[19, 20].

The experiment was tested with an informal pilot study[2] among colleagues before being put online at *redacted for anonymity*.

2.2 Measurements

The only two questions not directly relevant to the study were whether the subjects used a mobile device (as it changed the interface), and whether they had participated or seen someone participate in the study (as it could affect the memorisation). Considering the experiment's statistical power, we did not expect to be able to distinguish differences in demographic base performance. Thus, and out of a general concern over studies featuring irrelevant demographics questions, we decided to collect as little identifying data as possible — in accordance with local legislation.

All other recorded data relates to the answers provided to the questions asked during the study. For each of the three sets of pictures, we recorded:

- the index of the picture assigned to the subject (1 to 20);
- the description they gave for it;
- the list and order of each of the two sets of 10 images shown;
- the indices of the pictures they recognised, with a zero indicating that they chose "none";
- how much time they spent on each page.

We also recorded which group they were in for the A/B testing.

2.3 Image bank

During the design of the experiment, we conjectured that the recognisability and describability of the pictures would depend on what they depict. Distinguishing one human face from another is much more frequent for most people than distinguishing an abstract work of art from another. Moreover, the vocabulary to describe a human face is abundant, widely spread, and unambiguous when compared to the technical vocabulary required to describe a mountain or abstract works of art. The choice not to include human faces is discussed in Section 4.1.

We then used three series of 20 pictures with three themes: lions, mountains, and abstract shapes. The pictures were either (free of rights) pictures of lions or mountains, or they were abstract shapes randomly generated by the authors.

2.4 Subjects

Our study lasted from September 1st, 2020 to December 31st, 2020. We recruited subjects through John Krantz's Psychological Research on the Net index [14]. A total of 164 subjects participated to the study. All but two of them wrote their answers in English (with one French and one Spanish). The median time spent on the experiment was 213 seconds with a standard deviation of 169s — discounting users who took a noticeable break (between 20 minutes and 15 hours).

Before handling the full subject data, we removed from the study any subject who had not provided intelligible answers to the first part of the protocol (as a few adversarial subjects often participate). This eliminated a total of 13 answers, mostly corresponding to subjects who had skipped the questions, as well as a few who wrote descriptions such as "pee" or "po". This removal is not targeted towards the worst-performing subjects: of the 13 removed, 6 actually had perfect memorisation scores.

[2]The pilot study data is not included in our analyses as the protocols differ slightly.

A total of 4 subjects indicated that they had seen someone perform the experiment before or had performed it themselves. We chose to include those subjects as it is not evident in which direction the results would be affected if we excluded them: repetition could improve the performance, but could also increase the rate of false positive recognition. The performances of those subjects are similar to the other subjects and as such have little impact on the general results.

3 PRELIMINARY DATA ANALYSIS

3.1 Error types and frequencies

We interpret and analyse the answers to the image memorisation tests as a two-answer test rather than two independent tests. The subject succeeds if they get both answers right: if they manage to find the image they had been primed with earlier and click none on the other test. This allows us to compare their performance with an equivalent memory-less algorithm which chooses an image when not primed with one. Such a memory-less algorithm on independent tests would skew the results as the optimal strategy would be to always pick the "none" option, which is not realistic. On two-answer tests, such a memory-less algorithm would have at best a 5% success rate[3] (corresponding to picking exactly one image from each set of 20).

The following categorisation of errors was applied to each two-answer test:

- **False negative** (FN): the subject answered "none" to both queries;
- **False positive** (FP): the subject correctly found their picture but also recognised a second picture;
- **Single mistake** (SM): the subject correctly answered "none" to one query but chose the wrong image on the other query;
- **Complex mistake** (CM): the subject had a false positive on one query and a false negative on the other;
- **Double mistake** (DM): the subject chose the wrong image on one query and had a false positive on the other query.

Thus, a subject can accurately find their picture ('success') or have one of the previous four kinds of errors. Advanced statistical analyses are not relevant for the following data analysis for the following reasons:

- Comparing the success ratios to the null hypothesis (5% success rate) gives extremely high values (z-score > 40, corresponding to p-values < 10^{-350}).
- Comparing the success ratios between the different image series would not make statistical sense as the number of subjects is too low to perform a rigorous analysis of variance (ANOVA) between the image series. With the proportions measured, more than 1 000 subjects would have been needed to attain statistical significance.

The main subject accuracy performances are indicated in table 1 below:

	Correct	FP	FN	Mistakes (SM+CM+DM)
Lion	125 (83%)	12	6	8
Mountain	130 (86%)	8	10	3
Abstract	120 (79%)	12	9	10
1st image	128 (85%)	9	9	5
2nd image	127 (84%)	11	7	6

Table 1. Subject accuracy and error types for each series as well as for the first/second image shown (abstract always being third).

[3]This corresponds to an optimised memory-less algorithm, with a naive one having an 0.8% success rate.

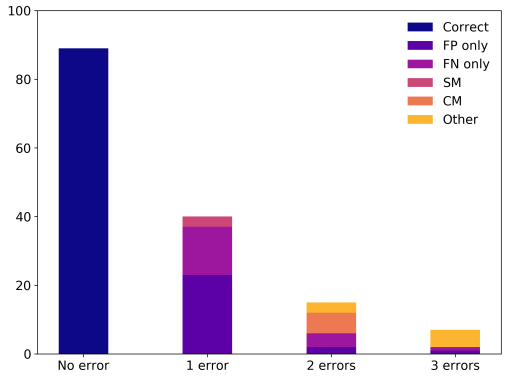

Fig. 1. Number and types of errors made by the subjects for the memorisation questions.

Figure 1 above shows the error types and the number of errors when trying to find their pictures. Among the 7 people who made three mistakes, one had only false positives and one had only false negatives. This might be due to misinterpreting the instructions, an effect we also observed and confirmed when running the pilot version of this experiment among colleagues.

3.2 Image describability and accuracy

To estimate image describability while minimising bias, two of the authors independently categorised the full list of descriptions subjects wrote about their assigned images, before comparing with the rest of the data. For each description, the assessors selected all images that could potentially fit. One of the assessors had the instruction to be strict in their estimates, and the other had the instruction to be lenient[4].

As can be inferred from Figure 2 (on the next page), some descriptions provide no information at all as they potentially correspond to all images. This is not always related to concision: although "lion" was the full description given by multiple subjects, some also provided other non-distinguishing descriptions like "the lion is the king of the jungle" or "70s, groovy, Brady Bunch, swirly, woman, colorful" (for the abstract image). We call these descriptions *trivial* if they can correspond to any image — for a given assessor. We discount them in some analyses to have a more rigorous interpretation of describability[5]. A few descriptions resulted in very different assessments, such as "the lion is focused on something", where the lenient assessor selected 17 images whereas the strict assessor selected no images. The accuracy of both categorisations (as the fraction of image selections that included the originally described image) is shown in Table 2 below.

Finally, in the goal of assessing the security of the images as potential visual secrets, we have the question of whether they can be accurately and unambiguously described — where a description fits a single image. Even with this noisy dataset, there is some evidence that certain unique elements get picked up by most subjects. Table 3 shows for each image series and assessor the number of unambiguous descriptions. It also shows how many of those descriptions considered unambiguous were in fact attributed to the wrong image (a low accuracy making the system more secure).

[4]In multiple cases, this meant that some parts of descriptions were ignored as potentially small mistakes. For example, 18% of subjects describe the abstract image as having blue among its main colours, although blue is very rare in the image set, and only twice makes up more than 15% of the image (when including many shades of blue).

[5]Our hypothesis and hope was that most users would provide ambiguous descriptions showing the difficulties in sharing their secret. Eliminating the worst performers imposes a stricter threshold on any subsequent result.

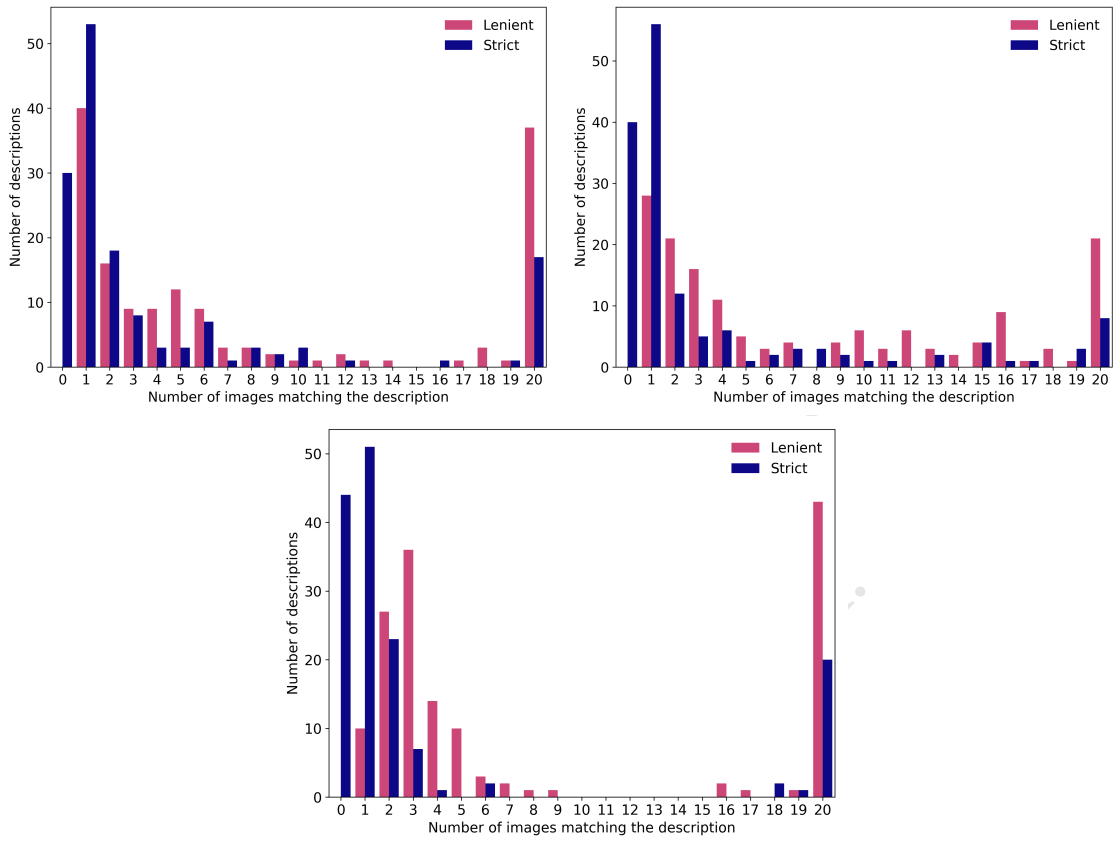

Fig. 2. Distribution of the number of selected images by the strict and the lenient assessor for the lion (top left), mountain (top right), and abstract (bottom) series.

	Assessor	Lion	Mountain	Abstract
All	Strict	59% (89)	55% (83)	50% (76)
descriptions	Lenient	92% (139)	78% (118)	76% (115)
Non-trivial	Strict	54% (72)	52% (75)	43% (56)
descriptions	Lenient	89% (102)	75% (97)	67% (72)

Table 2. Accuracy of the categorisations by the strict and the lenient assessors for the three series of images considering either all descriptions or only non-trivial ones (percentages with total number in parentheses).

	Assessor	Lion	Mountain	Abstract
Correctly	Strict	30	37	35
unambiguous	Lenient	27	21	7
Wrongly	Strict	17	12	16
unambiguous	Lenient	8	4	3
Unambiguous	Strict	64%	76%	69%
accuracy	Lenient	77%	81%	70%

Table 3. Number of unambiguous identifications by the strict and the lenient assessors for the three series of images. The proportion of correct identifications among unambiguous images is shown on the bottom lines.

4 DISCUSSION

4.1 Design choices

A central design choice was in the theme selection for the image series. We chose to avoid human faces for the following reasons:

- most humans[6] have specialised neural pathways that react specifically to faces [12, 21], which could create stronger reactions;
- this specialised facial processing's performance also depends on the age, ethnicity and gender of the face shown to the subject [29, 30];
- finally, languages tend to have specialised vocabulary to describe faces, which could improve the performance on describability.

For these reasons, we chose to use non-primate faces as they would not trigger human-specific responses [8]. As we still wanted to compare different types of image for describability, we settled on animal faces (lions) and natural scenes (mountains) [5]. For the third series, we wanted to have abstract images as they have the advantage of being easy to generate automatically — and as we conjectured that they would be hard to describe. We restricted the study to these three series to limit the time spent by subjects and the drop-out rate.

4.2 Limitations of the study

This study has one main limitation. We tested the subjects' memory only a few minutes after the initial stimulus. Although the writing the other descriptions to provide in between provided distractor tasks, the recognisability might still be influenced by short term memory effects.

A second potential limitation lies in the design as a web experiment where data is only stored if the user confirms at the end. First, this could limit the ecological validity compared to physically interacting with the pictures in a laboratory experiment with a controlled environment. Second, we could not measure the drop-out rate, and more importantly, the proportion of subjects who dropped out and redid the experiment. Two factors mitigate this. First, only a few subjects indicated having performed the experiment or seen it performed earlier. Second, other studies using the same source of subjects recorded a limited drop-out rate and next to no repeating subjects [4].

4.3 Future work on this dataset

This paper is late-breaking work with an empirical dataset that was compiled on December 31st, 2020. The data will be released upon publication, and the following questions can be further investigated in our data:

- How much variation is there in memorability between the images within each series? A cluster analysis could confirm the impact of pairs of very similar images.
- Similarly, as some images were apparently very easy to describe unambiguously, can clustering on the descriptions reveal any insight, and would this clustering match the one for the memorability?
- How is the memorability correlated with the time spent on the image description? How is it correlated (inversely) with the time spent on the description of other images (as they function as distractor tasks)?
- Does the mobile interface — used by 20% of subjects — have any impact on performance, compared to a full-sized computer screen?

[6] This sets aside people with certain neurodivergences and ones suffering from prosopagnosia, representing a non-negligible subset of the population [27, 28].

- How does describability vary by subject? Are some people better at describing everything unambiguously, or does it mostly depend on the images assigned to them?

4.4 Open problems

This work demonstrates that visual secrets are a viable security primitive. It also raises multiple questions about refinements and extensions:

- Could the image recognition process be fooled by images that are very similar, such as iterations on one basis made using generative adversarial networks [7]?
- Would visual secrets be viable with human faces, and how would one correct for demographic variation without knowing the subjects or users in advance?
- How unambiguous would the descriptions be if we allowed subjects to view the other pictures? What if we did so for a limited time, or only for a fraction of the image set?
- Using classical and quantum complexity formalisms, what constraints would allow non-shareable secrets?
- How sensitive to environmental conditions is the process? Our study happened *in situ*, but the images were presumably shown and recognised with identical screen settings. Would the performance be affected by the use of printed images or varying luminosity?
- As relying on sight alone could cause accessibility issues, would an auditory equivalent be viable? What kind of auditory stimulus would achieve the same recognisability, and would length be a hindrance (as sound is a more linear medium)?

5 CONCLUDING REMARKS

We have introduced a human security primitive called visual secrets, a kind of non-shareable secret that is pure information and do not depend on possessing an item. Its strength comes from two properties:

- the high recognisability of the pictures, with subjects have 80%+ accuracy;
- the difficulty of unambiguously describing the pictures. No assessor managed to get better than 81% accuracy on the 15-25% of descriptions which they thought were unambiguous.

This primitive allows new possibilities in terms of low-tech protocols that don't require complex sensors. The accuracy figures mean that the visual secrets could be used as is in specific contexts. For example, they could replace the identifying mark commonly used in verifiable voting systems inspired by Ron Rivest's ThreeBallot protocol [3, 24]. This would lower the probability of fraud detection from 33% to 28% per ballot, which would have no impact as it would be absorbed by the exponential behaviour when detecting fraud on multiple ballots.

Moreover, the accuracy could be amplified by the simultaneous use of multiple visual secrets (as the subjects have shown their ability to remember multiple images). For example, using 3 independent visual secrets means that the probability of fraud detection[7] is at least 99.1%. However, an adversary coercing the subject into describing their visual secrets would have at best an 1.5% chance of correctly identifying all the images.

Beyond voting protocols, we can also imagine visual secrets being use as part of authentication mechanisms or online communication protocols. We hope that this new primitive will inspire the development of systems with improved security and privacy in many different settings.

[7] In a voting setting, this assumes that the adversary changes the information attached to all the ballots. If they perform a single modification, the detection probability is at least 79%.

REFERENCES

[1] Roman Bednarik, Tomi Kinnunen, Andrei Mihaila, and Pasi Fränti. 2005. Eye-Movements as a Biometric. In *Image Analysis*, Heikki Kalviainen, Jussi Parkkinen, and Arto Kaarna (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 780–789.

[2] Enka Blanchard, Siargey Kachanovich, Ted Selker, and Florentin Waligorski. 2019. Reflexive Memory Authenticator: A Proposal for Effortless Renewable Biometrics. In *Emerging Technologies for Authorization and Authentication - Second International Workshop, ETAA 2019, Luxembourg City, Luxembourg, September 27, 2019, Proceedings (Lecture Notes in Computer Science, Vol. 11967)*, Andrea Saracino and Paolo Mori (Eds.). Springer, 104–121. https://doi.org/10.1007/978-3-030-39749-4_7

[3] Enka Blanchard and Ted Selker. 2020. Origami voting: a non-cryptographic approach to transparent ballot verification. In *5th Workshop on Advances in Secure Electronic Voting*.

[4] N. Blanchard, Clément Malaingre, and Ted Selker. 2018. Improving security and usability of passphrases with guided word choice. In *34th Annual Computer Security Applications Conference, ACSAC* (San Juan, PR, USA, 2018). 723–732. https://doi.org/10.1145/3274694.3274734

[5] Margaret M. Bradley and Peter J. Lang. 2015. Memory, emotion, and pupil diameter: Repetition of natural scenes. *Psychophysiology* 52, 9 (2015), 1186–1193.

[6] Bismita Choudhury, Patrick Then, Biju Issac, Valliappan Raman, and Manas Haldar. 2018. A Survey on Biometrics and Cancelable Biometrics Systems. *International Journal of Image and Graphics* 18 (2018). https://doi.org/10.1142/S0219467818500067

[7] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath. 2018. Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine* 35, 1 (2018), 53–65. https://doi.org/10.1109/MSP.2017.2765202

[8] Valérie Dufour, Michael Coleman, Ruth Campbell, Odile Petit, and Olivier Pascalis. 2004. On the species-specificity of face recognition in human adults. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition* (2004).

[9] Oded Goldreich and Yair Oren. 1994. Definitions and properties of zero-knowledge proof systems. *Journal of Cryptology* 7, 1 (1994), 1–32.

[10] Frank Haist, Arthur P Shimamura, and Larry R Squire. 1992. On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 4 (1992), 691.

[11] Wayne Jensen, Serban Gavrila, Vladimir Korolev, et al. 2003. *Picture Password: A Visual Login Technique for Mobile Devices.* Technical Report. National Institute of Standards and Technology.

[12] Mark H. Johnson. 2005. Subcortical face processing. *Nature Reviews Neuroscience* 6, 10 (2005), 766–774.

[13] Alexandros Kafkas and Daniela Montaldi. 2011. Recognition memory strength is predicted by pupillary responses at encoding while fixation patterns distinguish recollection from familiarity. *The Quarterly Journal of Experimental Psychology* 64, 10 (2011), 1971–1989.

[14] J. H. Krantz. 2019. Psychological research on the net. https://psych.hanover.edu/research/exponnet.html

[15] Sheng Li and Alex C. Kot. 2011. Attack using reconstructed fingerprint. In *IEEE International Workshop on Information Forensics and Security – WIFS.* IEEE, 1–6.

[16] Geoffrey R. Loftus. 1972. Eye fixations and recognition memory for pictures. *Cognitive Psychology* 3, 4 (1972), 525–551.

[17] Shane McCulley and Vassil Roussev. 2018. Latent Typing Biometrics in Online Collaboration Services. In *Proceedings of the 34th Annual Computer Security Applications Conference* (San Juan, PR, USA) *(ACSAC '18)*. ACM, New York, NY, USA, 66–76. https://doi.org/10.1145/3274694.3274754

[18] Marnix Naber, Stefan Frässle, Ueli Rutishauser, and Wolfgang Einhäuser. 2013. Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of vision* 13, 2 (2013), 11–11.

[19] Mizuho Nakajima and Yasushi Yamaguchi. 2002. Extended Visual Cryptography for Natural Images. In *The 10-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2002, WSCG 2002, University of West Bohemia, Campus Bory, Plzen-Bory, Czech Republic, February 4-8, 2002*. 303–310. http://wscg.zcu.cz/wscg2002/Papers_2002/A73.pdf

[20] Moni Naor and Adi Shamir. 1995. Visual cryptography. In *Advances in Cryptology — EUROCRYPT'94*, Alfredo De Santis (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 1–12.

[21] Charles A. Nelson. 2001. The development and neural bases of face recognition. *Infant and Child Development: An International Journal of Research and Practice* 10, 1-2 (2001), 3–18.

[22] Michael A Nielsen, Isaac L Chuang, and Isaac L Chuang. 2000. *Quantum Computation and Quantum Information.* Number 2. Cambridge University Press.

[23] David Noton and Lawrence Stark. 1971. Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research* 11, 9 (1971). https://doi.org/10.1016/0042-6989(71)90213-6

[24] Ronald L. Rivest and Warren D. Smith. 2007. Three voting protocols: ThreeBallot, VAV, and Twin, In Proceedings of the USENIX Workshop on Accurate Electronic Voting Technology (Boston, MA). *USENIX/ACCURATE Electronic Voting Technology – EVT*, 16–16. http://dl.acm.org/citation.cfm?id=1323111.1323127

[25] Roger N. Shepard. 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6 (02 1967), 156–163. https://doi.org/10.1016/S0022-5371(67)80067-7

[26] Ivo Sluganovic, Marc Roeschlin, Kasper B. Rasmussen, and Ivan Martinovic. 2016. Using Reflexive Eye Movements for Fast Challenge-Response Authentication. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (Vienna, Austria) *(CCS '16)*. ACM, New York, NY, USA, 1056–1067. https://doi.org/10.1145/2976749.2978311

[27] Tirta Susilo and Bradley Duchaine. 2013. Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology* 23, 3 (2013), 423 – 429. https://doi.org/10.1016/j.conb.2012.12.011 Social and emotional neuroscience.

[28] Lucina Q Uddin, Mari S Davies, Ashley A Scott, Eran Zaidel, Susan Y Bookheimer, Marco Iacoboni, and Mirella Dapretto. 2008. Neural basis of self and other representation in autism: an FMRI study of self-face recognition. *PloS one* 3, 10 (2008), e3526.

[29] Holger Wiese. 2012. The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology* 89, 1 (2012), 137 – 147. https://doi.org/10.1016/j.biopsycho.2011.10.002

[30] Nicole Wolff, Kathleen Kemter, Stefan R Schweinberger, and Holger Wiese. 2014. What drives social in-group biases in face recognition memory? ERP evidence from the own-gender bias. *Social Cognitive and Affective Neuroscience* 9, 5 (2014), 580–590.

6 APPENDIX: IMAGE GALLERY

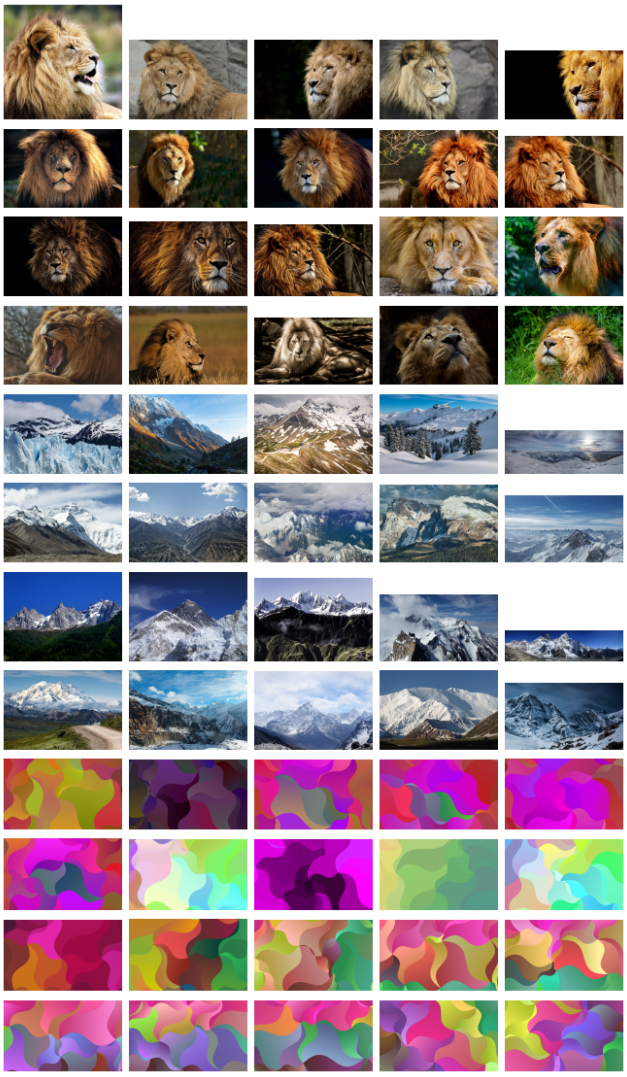

Fig. 3. The three image series used for the experiment.