Visual Secrets : A recognition-based security primitive and its use for boardroom voting

Enka Blanchard, Sébastien Bouchard, Ted Selker

▶ **To cite this version:**

Enka Blanchard, Sébastien Bouchard, Ted Selker. Visual Secrets : A recognition-based security primitive and its use for boardroom voting. 2022. hal-03133412v2

HAL Id: hal-03133412
https://hal.science/hal-03133412v2

Preprint submitted on 16 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Visual Secrets : A recognition-based security primitive and its use for boardroom voting

Enka Blanchard
CNRS
LAMIH, Université Polytechnique Hauts-de-France, Valenciennes
Center for Internet and Society, Paris, France

Sébastien Bouchard
Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

Ted Selker,
Cyber Defense Lab
University of Maryland, Baltimore County (UMBC), USA

Abstract. This paper presents and evaluates a new security primitive in the form of non-transferable "visual secrets". We show how they can be used in the design of voting systems. More specifically, we introduce a receipt-free low-tech visually verifiable boardroom voting system which is built for simplicity and can serve as a teaching tool to introduce people to verifiable voting.
Visual secrets rely on the pre-semantic treatment of images in the human brain. After being shown an image for a limited time, users can recognise it in a larger set (79% recognition compared to an expected baseline of 5%). However, they cannot reliably communicate to someone else exactly how to do so (whether voluntarily or through coercion). Indeed, no assessor managed to accurately identify the images from the description given by the subjects.
We then introduce a boardroom voting system based on this primitive. The voter receives a ballot consisting of a single picture, votes by folding it horizontally or vertically and casts it. If the voter is coerced into describing their ballot, several ballots are likely to correspond to the description. When all ballots are revealed, the voter can check with a glance that their ballot is present and folded correctly. This gives them the opportunity to detect error or fraud, although they cannot prove that fraud happened (the limited dispute resolution mechanism mainly focuses on incentives). The design makes use of textured paper to provide both accessibility for the blind and improved usability for all users.

Keywords: Usable security, Boardroom voting, Verifiability, User studies, Cognitive psychology

1 Introduction : defining visual secrets

Despite many advances in verifiable voting over the last 20 years, two problems are nearly as relevant today as they were then. First, there is limited under-

standing by the public of both how verification works, and why voting systems should be verifiable (irrespective of the cost-effectiveness) [9]. Second, the usability costs remain high, both for end-users and administrators, limiting the number of users who verify their votes [14, 39]. For example, long vote-codes remain prevalent, and users can get confused as to their purposes [9]. We initially sought to improve usability by simplifying those codes or finding equivalent presentations, and instead found a new security primitive that could have multiple applications. One such application is the central component of a simple verifiable voting system meant to introduce users to the concept of verifiable voting.

The secrets employed in usable security often correspond to "something you are, know or have". However, most such secrets are shareable : one can give their home keys to a friend, be coerced into revealing passwords, or even have their biometrics such as fingerprints stolen [26]. One natural question is then to ask whether it is possible for humans to have (useful) secrets that cannot be shared ? In a formal way, the answer seems to be no, but if we set reasonable constraints, some tentative solutions can be found. One option could be a behavioural biometric, which is possible in the abstract case, but multiple problems exist with those, from high error rates to biometric identity theft [28, 10]. Moreover, this type of secret can require complex apparatus to measure.

A second lead is then to use specialised human cognitive functions. There have been some measure of success in creating unshareable secrets in [7], as subjects have no conscious recollection of them, but the training is time-intensive. One cognitive 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 [35, 22]. This has already been used as a source of security primitives, for example with authentication in the case of visual passwords [18], as well as with various biometric methods [3, 37]. Most importantly in our case, a significant aspect of this image recognition happens in a pre-semantic and pre-cognitive fashion, requiring no conscious effort, thanks to specialised neural pathways in multiple areas of the brain [29, 22]. This is related to the difference between recognition and recall [17]. 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.

Our approach takes inspiration from both this cognitive science research and concepts from zero-knowledge proofs [15]. This pre-semantic treatment is used as a source of secrets that are recognisable but not shareable, and we call the resulting primitive a *visual secret* (which are not related to visual cryptography [31, 30]). 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.

These constraints are especially appropriate in our case, as the first proposed application of visual secrets concerns verifiable voting in a boardroom setting. This corresponds to a small group of participants — e.g., jury members — having to quickly vote on an issue, generally between two possibilities. In practice,

such votes are often held informally by writing an answer on a piece of paper. The assumption that this is secure often relies on the high cost if an attack is discovered compared to the generally low stakes and some attacks have been proposed against such systems [5]. For example, an attack that is doable with little training consists in depositing two papers and drawing back one from the bag (which statistically gives a single vote advantage to one candidate. Quite naturally, a variety of options have been proposed that seek to improve on this naive method [16, 1, 27].

The central idea behind our application is to have a visual secret on each ballot — but no receipt except for the memory of said visual secret. This allows each voter to check that their ballot is present and counted correctly, but the receipt-freeness [21] prevents them from proving to someone else that they voted a certain way. For this application, two metrics are crucial : a high short-term recognisability (to find one's own ballot) and a low describability (several ballots could correspond to one's description of one's own ballot).

Over the next five sections, we first relate the results of a usability study on 151 subjects that demonstrates that visual secrets have both high recognisability and low describability. We then introduce the ballot design and its security analysis. Finally we discuss the limitations of our work and conclude with future research leads.

2 Empirical study

The goal of the study was to test the viability of visual secrets as a security primitive. Subjects were shown three pictures and had to describe them, before having to find their initial pictures among sets of 20 similar pictures in random order. Additional information on the protocol, the results of the study on 151 subjects and statistical analyses (including non-significant tests) are below. In the goal of transparency, all the corresponding data in this section has already been put online and is publicly available as an attachment to the Hal version of this document.

2.1 Image choice

As we conjectured that the recognisability and describability of the pictures would depend on what they depict, three different image series were included. While people are in general much better at recognising human faces, we chose to avoid them as:

- performance varies highly depending on the age, ethnicity, and gender of the face shown to the subject [40, 41];
- subjects might also have stronger emotional reaction to faces [32, 19];
- many languages have specialised vocabulary to describe faces — which is more widely spread than the technical vocabulary required to describe a mountain — which could improve the performance on describability;

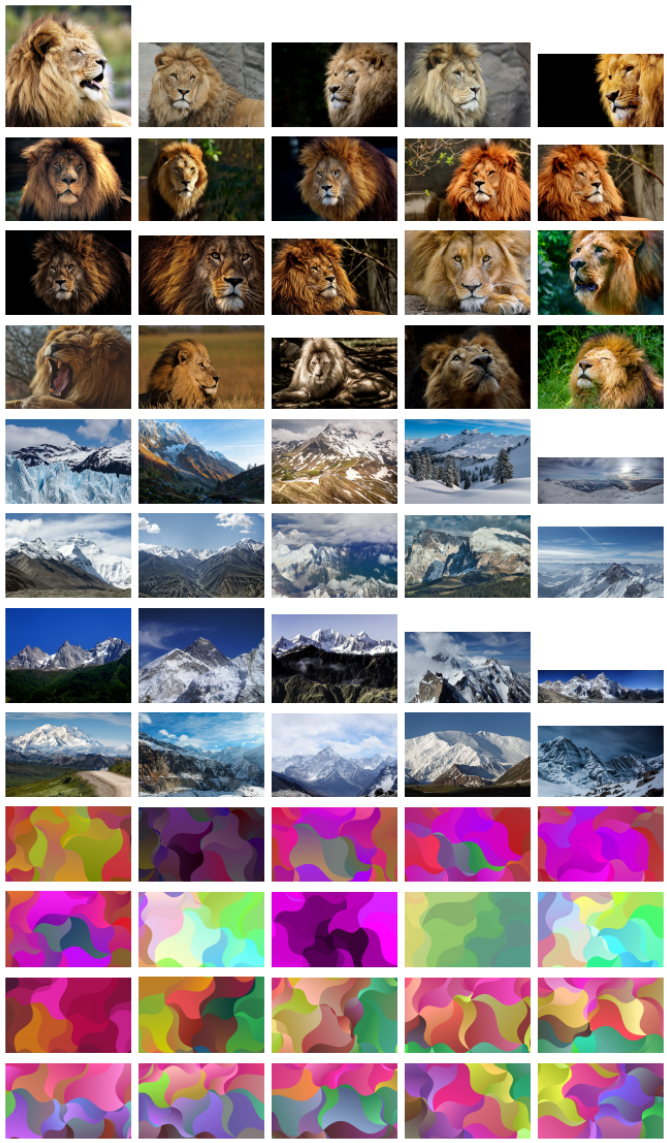

Fig. 1. The three image series used for the experiment.

For these reasons, we chose to use non-primate faces as they would not trigger human-specific responses [13]. For the first two series, we settled on public domain images of animal faces (lions) and natural scenes (mountains) [8]. 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 harder to recognise and especially to describe. We restricted the study to

these three series to limit the time spent by subjects and the drop-out rate. Figure 1 shows the three image series.

2.2 Experimental design

Figure 2 summarises the experimental design. The *Basic information* included an 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 (4 answered yes and we chose to include them as their performances are similar to the other subjects and as such have little impact on the general results). The *Describe* pages showed a picture and asked to describe the picture in at most 10 words to try to make it identifiable among similar images.

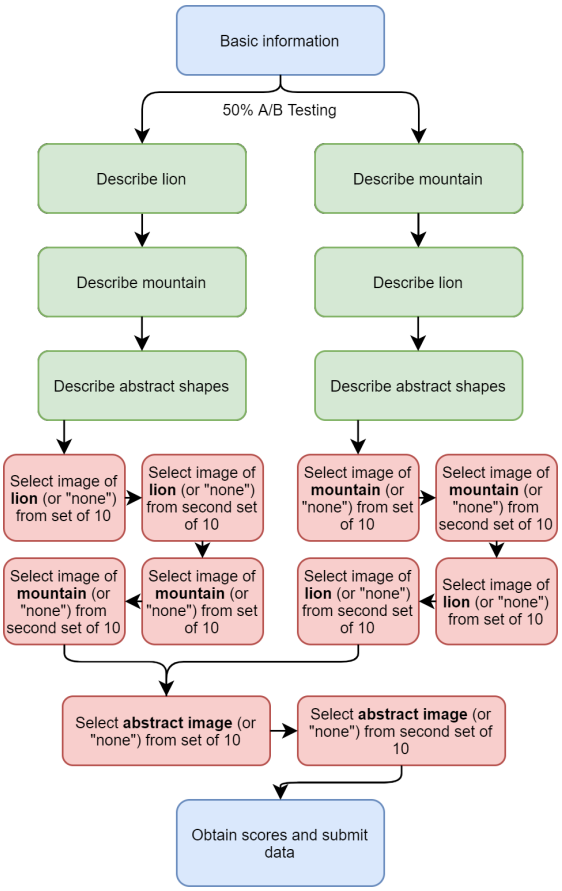

Fig. 2. Diagram of the experimental protocol.

A/B testing was used to randomly assign the order of the first two image series (lion and mountain), with the image recognition order being the same as

the presentation order. The third series was always shown last. The tasks of writing the second and third images descriptions then served as distractor tasks in order to limit the effect of short-term memory. The A/B testing allowed us to measure and compensate the effects of order and delay on recognition, with no significant effect observed.

The only two questions not directly relevant to the study were whether the subjects used a mobile device (as it changed the image layout), 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.

The experiment was first tested with an informal pilot study among colleagues before being put online at `http~://koliaza.com/visualsecrets/`. The pilot study data is not included in the analyses as the protocols differ slightly.

2.3 Subjects

We recruited 164 volunteers through John Krantz's Psychological Research on the Net index [23]. No protected demographics were targeted, all the participants were informed of their rights and could quit at any point they wanted, no deception was used, and the data was only collected if they confirmed at the end. Besides the main language spoken (English with two exceptions), no personal or identifying information was collected. We eliminated subjects who had not provided intelligible answers when asked to describe pictures, leaving 151 subjects. This removal is not targeted towards the worst-performing subjects : of the 13 removed, 6 actually had perfect memorisation scores.

2.4 Recognisability

We chose to 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 both find the image that had been shown earlier and click "none" on the other test. Otherwise they had false negatives (FN) if they answered "none" twice, false positives (FP) if they recognised a second picture, and other mistakes (OM) otherwise.

This allows us to compare their performance with an equivalent memory-less algorithm which chooses an image when not primed with one. Without the two-answer approach, 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 neither useful nor realistic. On two-answer tests, the memory-less algorithm would have at best a 5% success rate (corresponding to picking one image from each set of 20). This corresponds to an optimised memory-less algorithm, with a naive one having an 0.8% success rate.

As shown in Table 1, subjects could recognise their pictures with high reliability (79% to 86% rate). When compared to a null hypothesis of 5% (for optimised random choice), this is highly significant (z-scores >40 for all series, corresponding to p-values $< 10^{-350}$). However, as the error rates were lower than expected, comparing them between the series is not within the statistical power of the experiment, and would have required more than 1 000 subjects assuming similar values (hence, no comparisons are made between the series as they would all have p>0.05).

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

Table 1. Subject accuracy and error types for each series as well as for the first/second image shown (abstract always being third). No significant effect was seen between pictures within series.

Figure 3 shows the error types and the number of errors when trying to find a previously seen picture. 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 also observed at much higher rates when running the pilot version of this experiment among colleagues (we clarified the instructions to address this after the pilot).

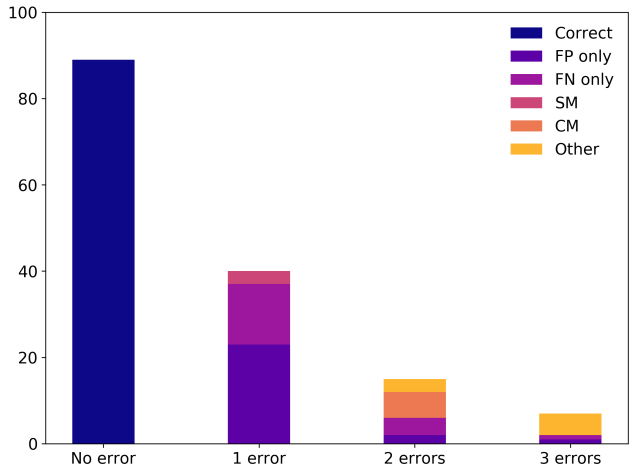

Fig. 3. Number and types of errors made by the subjects for the memorisation questions. Other mistakes are split into SM (FN plus the wrong image), CM (FN + FP), and others.

We computed the error rate independently for each image, which gave no statistical result. As could be expected, all the images have statistically similar rates of being chosen incorrectly, and the ones with a higher proportional error rate correspond to the ones that were shown the least — thus increasing

the proportional variance (in practice, no image was misidentified more than 3 times). We also looked at whether the recognition errors tend to form clusters of related images, but the corresponding graphs are too sparse to hypothesise the presence of any definite cluster.

2.5 Image describability

Describability To estimate image describability while minimising bias, two of the authors independently categorised the full list of descriptions subjects wrote about their assigned images. For each description, the assessors selected all images that could potentially fit — without knowing what the correct answer was. One of the assessors had the instruction to be strict in their estimates, and the other had the instruction to be lenient. In multiple cases, this meant that some parts of descriptions were ignored as potentially small mistakes. For example, although blue is very rare in the image set, 18% of subjects describe the abstract image as having blue among its main colours. This provided upper and lower bounds for multiple metrics.

Figure 4 shows the number of selected images by each assessor, which shows that 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 *trivial* if the assessor selected all images. We discounted them in some analyses to have a more rigorous interpretation of describability (eliminating the worst performers imposes a stricter threshold on any subsequent result). The accuracy of both categorisations (as the fraction of image selections that included the originally described image) is shown in Table 2. 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.

	Assessor	Lion	Mountain	Abstract
All descriptions	Strict	59% (89)	55% (83)	50% (76)
	Lenient	85% (129)	78% (118)	76% (115)
Non-trivial descriptions	Strict	54% (72)	52% (75)	43% (56)
	Lenient	81% (92)	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).

In the goal of assessing the security of the images as potential visual secrets, one question is crucial : can they be accurately and unambiguously described, or in other words, does a description fits a single image ? Even with this noisy dataset, there is evidence that certain unique elements get picked up by most

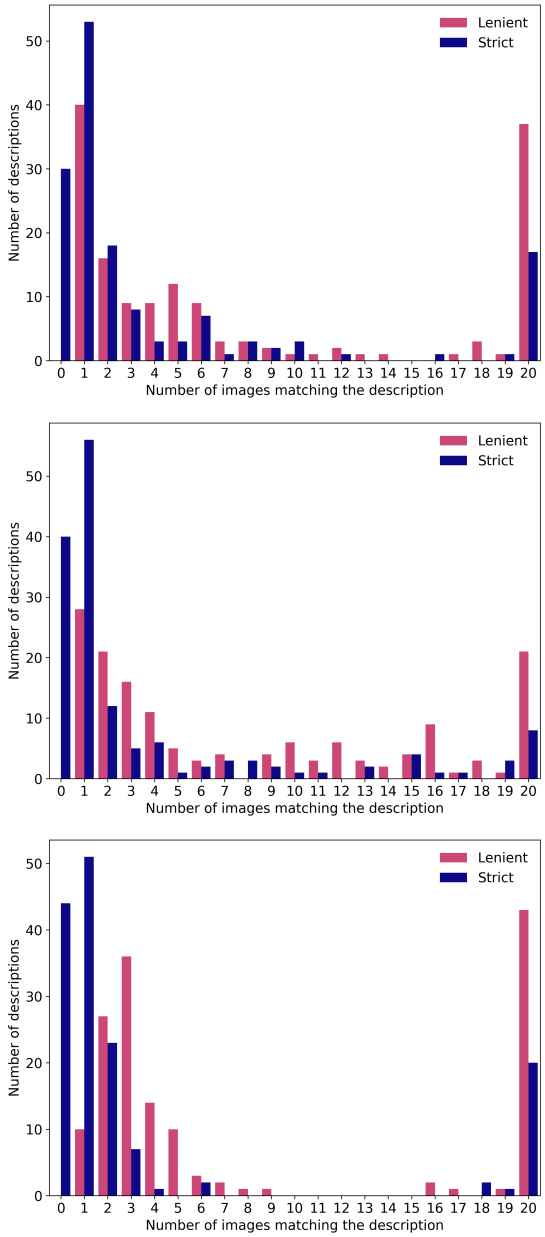

Fig. 4. Distribution of the number of selected images by the strict and the lenient assessor for the lion (top), mountain (middle) and abstract (bottom) series.

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).

	Assessor	Lion	Mountain	Abstract
Correctly unambiguous	Strict	36	40	35
	Lenient	32	23	7
Wrongly unambiguous	Strict	17	16	16
	Lenient	8	5	3
Unambiguous accuracy	Strict	68%	71%	69%
	Lenient	80%	82%	70%

Table 3. Number of unambiguous identifications by each assessor for the three series. The proportion of correct identifications among unambiguous images is shown below.

As shown above, the accuracy is at most 82%, even when restricting ourselves to when the assessors were sure of their choice. We've thus established that visual secrets are close to our objectives. They are highly recognisable (79-86%) while being poorly describable. A coercer trying to obtain the secret would then have succeeded in at most 26% of cases, with an additional 8% of cases where they would have been (wrongly) sure that they had found the correct secret.

Image clusters One last test we tried was to compute clusters of images likely to be all selected by either assessor when one of them is described, therefore indicating images for which it is unlikely that the subjects provide a description distinguishing one from another. A first step was to find whether certain images often get confusing descriptions. For example, when faced with descriptions of abstract image #2, the lenient assessor considered that image #6 was potentially the one being described twice as frequently as image #2. Table 4 shows the number of images for which the descriptions generally point to a different image.

	Most probable		Among probable		Not probable	
	Lenient	Strict	Lenient	Strict	Lenient	Strict
Lion	10	12	6	3	4	5
Mountain	8	10	6	4	6	6
Abstract	6	8	6	5	8	7

Table 4. For each series, this table shows the number of images for which the descriptions tended to correspond to other images more than their original image. For each image, if its descriptions most frequently point to itself being selected; it is counted in "Most probable". It is counted in "Among probable" if there are other images as frequently assigned to its descriptions, and "Not probable" otherwise.

We also computed a full graph for each (assessor, series) pair. Despite the limited number of descriptions, none of the graphs are sparse (with 145 to 304 edges out of a maximum of 484), making them hard to interpret. Figure 5 shows one such graph (the most legible one, having the least number of edges). The noisy nature of the dataset limits the interest in deleting the low-weight edges.

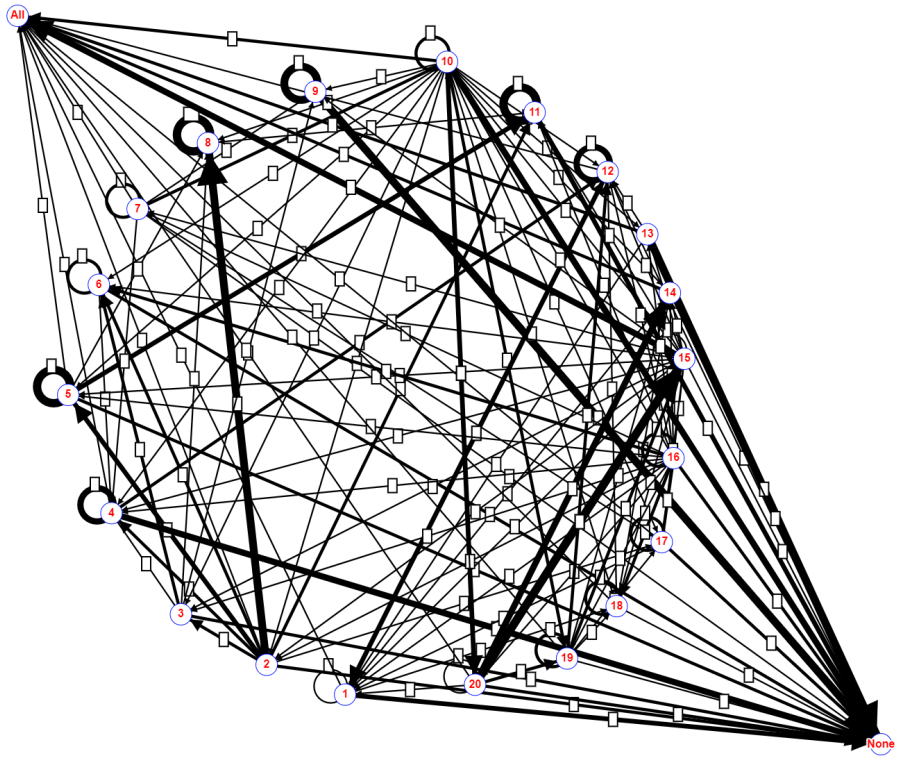

Fig. 5. Graph of the abstract shape descriptions by the strict assessor. Edge thickness is proportional to how often a description for i was interpreted as one for j, with two added nodes ("None" and "All") to merge trivial descriptions and make the graph more legible. In practice, we can see that the main outgoing edge of node 8 is a loop on itself, meaning that it was generally well categorised. Conversely, descriptions of image 2 were mostly badly categorised, and often interpreted as corresponding to image 8 (hence the thick edge from 2 to 8).

Despite the difficulties in handling this noisy quantitative data, some effects are apparent, and confirmed by qualitative feedback from both assessors. For example, the lenient assessor managed to unambiguously recognise every description of the #20 lion picture (and the strict assessor performed almost as well). The peculiarity is that the unique features pointed out by the subjects varied a lot, e.g. the sleeping/closed-eyed lion or the yellow flowers. This image was not identified as particularly describable when creating the dataset, so care should be taken when creating new datasets as empirical validation is crucial.

3 Visually Verifiable Ballots (VVB)

We now describe a first application of visual secrets in the form of a low-tech — in our case, paper — voting system appropriate for boardroom elections. VVB

has two central objectives. The first is providing a low-tech system that is not subject to the attacks mentioned in Section 3.1 below. The second is to provide a cheap[1] teaching tool that is easy to use and can introduce users to the concepts of verifiable voting (before moving on to more secure and complex systems such as Belenios [11]).

3.1 Threat model

One issue with the common kind of boardroom voting is that it is hard to estimate the prevalence of errors and fraud, which can be high in corporate settings [2]. The traditional "handwritten vote on a piece of paper in a bag" method has multiple flaws. It makes it possible to leave identifying marks or simply to recognise someone's writing, and an adversary with some limited skill in prestidigitation could steal a ballot or two from the bag when casting their own ballots (which can be prevented by having a transparent urn, but this is rarely done). One advantage, however, is they are generally in-person boardroom election. This means that the cost of being caught is high and that the first priority of the adversary should be to remain covert and as unsuspected as possible. If caught, they could be banned and the election would proceed without them with little delay, as opposed to a large-scale election.

As done in [5], we assume that the adversary can have some accomplices, as well as some skills in prestidigitation, allowing them to dissimulate and manipulate paper ballots in a discreet way. They can also try to coerce other voters but not cannot do so publicly — hence the interest of receipt-freeness for potentially coercible voters. Moreover, we assume that they cannot convince or coerce the majority of voters into doing their bidding. As many elections use a simple majority, this would make many objectives irrelevant. It also opens multiple new avenues of attack depending on the degree of control the adversary has on the other voters. We also assume that they do not have access to high-tech systems (such as hidden cameras in the room), except for common items (such as smartphones). We consider an adversary with the following four potential objectives :

1. directly change the outcome of the election ;
2. find out how other people voted ;
3. coerce others (or pay them) into voting for a designated candidate ;
4. cast doubt on the outcome (discreditation attack).

3.2 Ballot design

Visually Verifiable Ballots look and feel like square cards (an example is shown on Figure 6). Just like cards, one side is left blank — or with a regular symmetri-

[1] Visually Verifiable Ballots could be made available as packs of 20 to 50 ballots, wrapped and sealed like a playing card deck. Initial estimates show that the cost to manufacture such packs should be between $0.50 and $1 per pack, using commercial printing services.

cal pattern — and the other has the relevant information. The visual information is minimal, as it consists of two elements :

- A picture from a common set of visual secrets, covering the whole card ;
- Two orthogonal lines crossing the picture, labelled “Vote 1” and “Vote 2”.

This visual information is complemented by tactile information in the form of texture — bumps — present on both ends of each line, with one bump for the first and two for the second[2]. This has two objectives. First, it allows voters to keep track of the ballot’s orientation without seeing the lines. Second, it also makes it possible for visually impaired people to vote without requiring a different voting system (although they would not be able to verify). Care should be taken when applying the tactile patterns (for example, using thick ink instead of mechanical embossing) to avoid the bumps being noticeable from the other side — as well as to avoid transparency issues.

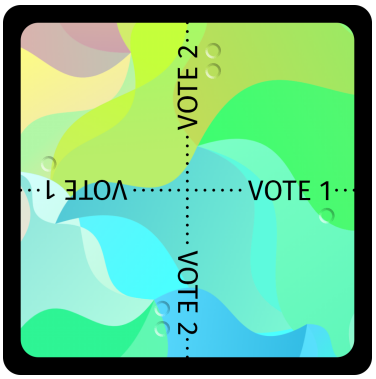

Fig. 6. Example of a Visually Verifiable Ballot. Folding along a line correspond to voting for that option. The embossed bumps represented under the folding lines help voters keep track of which line is which.

3.3 Protocol

The protocol goes as follows :

1. The vote organiser opens a new pack of ballots in front of all voters ;
2. One ballot is distributed face down to each voter ;
3. Each voter lifts up their ballot to look at the image and memorise it ;
4. Each voter rotates their ballot a few times, while keeping track of its orientation using the bumps ;
5. Each voter folds their ballot along the line of their choice to select “Vote 1” or “Vote 2” to be on the inside fold, without marking or modifying their ballot in any other way ;

[2] This coincidentally corresponds to International Braille for A/B, which is only of limited benefit, as Braille teaching becomes increasingly rare [33].

6. The voters cast their ballots in a ballot box or a bag ;
7. The ballot box is upturned and all the ballots are unfolded on a table in front of all the voters' eyes ;
8. The vote organiser tallies the votes orally while the voters check that the ballot featuring their assigned picture are present with the correct fold ;
9. If a voter sees their ballot folded the wrong way or cannot find their ballot, they announce as much without giving any additional information ;
10. The vote organiser announces the result and the vote is over unless someone challenges the result.

4 Security considerations on VVB

This section shows how VVB addresses the four attacks mentioned in Section 3.1. We focus on abstract visual secrets for the analyses, as they are the worst-performing in terms of recognisability and thus give a lower-bound on the expected performance of the other image series (as small variations in describability have limited impact). The next two subsections assume that the voter is not an accomplice or being coerced by the adversary — which will be covered in the following subsections.

4.1 Changing the outcome

Assuming the ballots are publicly counted, an adversary wanting to change the outcome must manage to do at least one of the followings :

- add ballots ;
- remove ballots ;
- replace ballots.

As there are at most a few dozen voters and the vote counting is done in plain view, adding or removing ballots at any point would introduce discrepancies. The goal is then to replace ballots without voters noticing.

As long as voters check their ballots, the only way to replace ballots without being noticed is for a single ballot to be recognised by multiple voters. This could be achieved thanks to a complex attack with a specially made pack of visual secrets, the ability to distribute those selectively, then to remove them from the ballot box before adding other ballots. Let's consider that the adversary does not resort to such an attack and the voters' recall ability is standard according to our experiment. Then changing the fold on a single ballot or replacing it by a different ballot would get noticed with at least 79% probability (however, folding a ballot twice would in itself leave marks). Changing multiple ballots without anyone noticing would then have an exponentially smaller probability (in the number of modified ballots).

However, there might also be false positives. Assume that there are twenty voters with a split vote[3]. We would then expect 2 or 3 voters to be mistaken

[3] If the vote was strongly in favour of the adversary's choice, there would be no need to cheat, and if it was strongly in their disfavour, the following analyses would give even higher chances of getting caught).

or confused about their ballots (but probably not sure of themselves). Having 3 voters notice errors (and be sure of themselves) would then be a strong indicator of malfeasance, and would be a probable outcome even if only a few ballots are replaced.

4.2 Finding out how other people voted

If the adversary can see the front of the ballot (including the visual secret), they could remember it and check during the tally. However, each ballot is distributed face down, and each voter should be careful to prevent the front of their ballot from being seen by anyone else. In any case, viewing the front of ballot of another voter is mostly doable by adversaries sitting next to them. Let us now suppose that the adversary can only see the back of the ballot and not the front until all ballots are revealed during step 7.

If an adversary manages to find out how someone voted, they necessarily learn this information at the earliest when the voter folds their ballot, that is, during step 5 of the protocol. We consider steps 5 to 9, and list the means implemented in the protocol to prevent any adversary from learning how other people voted during these steps.

We start with step 5. Finding out how another person voted is equivalent to finding out whether the fold is vertical or horizontal (with respect to their ballot). Either the adversary knows the initial orientation (before the ballot is given to the voter) or they do not. If they do, in order to use this initial information, they should also keep track of the rotations, which requires visual acuity and would be hindered by some voter actions (e.g., rotating their ballot under their hands). Otherwise, figuring out the orientation of the ballot is hindered by the symmetry of the back of the ballot. Having an identifying mark on the back of the ballot that would be visible from a distance would allow to overcome this symmetry but would be noticeable — if not necessarily noticed — by both the voter and other people in the room.

During step 6, when the ballots are cast, the fronts of the ballots are at least mostly hidden thanks to the fold performed during step 5. The difficulty to see other people's ballot introduced by the fold of step 5 could be strengthened as discussed in Section 5.3 by adding a thicker black border on the edges.

We now consider step 7, when all ballots are revealed. Let us first assume that the adversary has prior knowledge of the set of visual secrets being used. A first option would then be to keep track of which ballots were given to which voters during distribution, after which it is trivial to know how everyone voted. However, this would require having two identical unopened packs of visual secrets, and for the adversary to be able to impose using one of those packs. Moreover, the adversary should then control the distribution of the pack of visual secrets. Having someone shuffle the pack beforehand would then address this issue as long as the person shuffling is not an accomplice. Attacks of this kind would still be possible, but would require technical ability, equipment, and the presence of at least one accomplice.

Let us now assume that the visual secrets are not known beforehand. An adversary — especially one who distributes ballots — could take the opportunity during the distribution to make an identifying mark on the back of certain ballots, observe which voters got them, and then look for those marks during the tallying. As above, a mark visible from a distance would have a high chance of being caught during the counting, so the mark should both be discreet and yet easily findable during the tallying — which seems difficult to arrange. A related risk would be for voters to fold ballots incorrectly in a noticeable fashion, but the folding lines and the number of voters limit this attack's feasibility.

During the steps 8 and 9 of the protocol, each voter checks that the ballot featuring their assigned picture is present with the correct fold. No matter what they observe, unless they broke protocol in the previous steps, voters cannot prove that a ballot belongs to them (hence the receipt-freeness). As stated in step 9 of the protocol, voters also do not have to indicate which ballot was theirs to claim that there was an error, and can hence do this without revealing how they voted. The impossibility of proving malfeasance introduces the possibility of making false claims, an attack which is covered in subsection 4.4.

4.3 Coercion and vote-selling

The most common way of coercing people (or paying them) into voting for a designated option is to have a way to find out how they voted, with their collaboration. This can happen in two distinct ways: they could directly show their ballot to someone, or they could create and share a receipt.

The easiest option is to have a voter discreetly show their ballot during the voting period (or take a picture with their phone), as proof that this is their visual secret. However, voters should only have time to get their ballot, peek at it discreetly while keeping it face-down, think briefly, rotate and fold it, and then cast it. As each voter can see the other voters, showing one's ballot to someone else — except maybe one's immediate neighbour — would be risky and the culprit could get caught.

If voters can't show ballots during the casting phase, they could try to describe it (by speech or writing) — which is equivalent to creating a (possibly intangible) receipt. Once the ballots are public, the receipt-freeness means that nothing prevents the voter from lying and describing a different ballot. If they try to describe their ballot before the count, they must do so discreetly and succinctly as they are in the same room with limited time — also reducing the scalability of such process. Let's assume once again that the adversary has no prior knowledge of the visual secrets. Even in the best case, no assessor managed to identify accurately and without ambiguity more than 26% of a series of visual secrets from their descriptions, limiting the interest of such a method[4]. Moreover, the

[4] There is a natural difference between interactive and passive protocols, and our experiment was based on a passive protocol where participants provided description without being prompted with questions. However, as we assume that the adversary has no prior information on the visual secrets used, we do not see that there would be any advantage following an interactive question/answer protocol.

subjects were describing their pictures while being able to observe them, as opposed to describing from memory, which would further reduce performance. This is where visual secrets have an advantage over other kinds of secrets such as random codes[5]. It would be pretty easy in most settings to write down a few characters from one's code or even the whole code if it were short (this could be done while the bag is being opened or any other distraction). If we go back on our hypothesis and assume that the adversary has prior knowledge of the set of visual secrets, they could also show the ballots to their coercion victims and agree on a code so the latter can easily indicate which ballot they had before the tallying. This is why we recommend using sealed randomly-generated packs.

If we put aside finding out how someone voted, another standard option for a coercer is to vote in someone's stead and prevent them from reporting it — or to give them an already filled-out ballot, as with chain voting [20]. Voting directly in someone else's stead is impossible in a boardroom. However, if a bag or an opaque box is used as a ballot box, one attack could hypothetically be achieved. It would require the coerced voter to simulate dropping their ballot in the bag (while keeping it in their hand). The adversary could then add two ballots simultaneously without alerting anyone. The difficulty with this attack is that it requires some sleight-of-hand ability in both the adversary and the coerced voter, and is irrelevant if the ballot box is transparent.

4.4 Discrediting the election

A small group of adversaries could falsely claim that their ballots were modified to create a discreditation attack — presumably cancelling the vote. This is the case with all verifiable voting systems in which voters can only detect but not prove the existence of malfeasance. As such, our system is vulnerable to this kind of attack. However, as all voters are in the same room, they can easily start a new election, potentially changing the people in charge of handling the election or even using a different system — ideally more secure. This kind of attack would then mostly delay the election by a small amount while bringing unwanted attention to the group of adversaries. Moreover, if the election is to be decided by a simple majority and a single voter reports irregularities while the margin is of at least a few votes, the voters could decide (in advance) to accept the result.

5 Discussion

5.1 Limitations of the experimental study

This experimental study has one main limitation. It tested the subjects' memory only a few minutes after the initial stimulus. Although writing the other descriptions in between provided a distractor task, the recognisability might still

[5] Designing alphanumeric codes that are highly recognisable (within a set of similar codes) while being hard to write down (even in part) seems a difficult endeavour.

be influenced by short-term memory effects. However, the impact of time spent between the memorisation and the recognition had no evident (or statistically significant) effect. There was a tendency to have slightly lower recognition when considering subjects who had a delay between 3 and 15 minutes, counteracted by the fact that many of the slowest subjects had perfect scores.

A second limitation is that the subjects were recruited from one of the major sites that indexes psychological studies online (with around 500 studies hosted each year), which could have created a recruitment bias. However, previous studies have shown that the participants from this subject pool tend to give higher-quality data than users of Mechanical Turk, and that they cover a wide range of demographics, albeit with a bias in favour of college-age respondents [24].

A third potential limitation lies in the study's 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 [6].

Finally, there is the question of whether participants could have described the images more precisely if they had been more informed or better motivated. Responses to online open-ended questions are often short (although we tried to limit this by having a large text-box) [38]. Beyond specialists (zoologists for the lion images, or colour theorists) who already have the expertise to describe the images, non-expert participants could also potentially learn how to describe the images more accurately, which would be a weakness if the system is used repeatedly with similar image banks. One way to test the limits of this would be to ask participants to describe an image in an unambiguous way when having access to the full set of pictures (or half the set), as proposed in Section 6.1.

5.2 Considerations on using VVB in practice

From what we have shown, VVB should be reasonably secure when used with small groups of voters who can see each other and who vote rapidly, thus benefiting from the potential short-term memory advantage. This can be in low-importance votes (for example, promotion committees), but not only, with one particular application coming to mind being juries in legal cases. VVB is particularly suited to this situation for multiple reasons : the set of voters is small and they do not know or trust each other ; they might want to organise multiple votes, so each one should take little time ; they are not expected to be skilled at sleight-of-hand. As juries are often decided shortly before a case, it is also hard to train a jury member to perform certain attacks.

One limitation of VVB is that the version shown above does not allow voters to anonymously abstain but only to vote between two possibilities. One variant

shown in Section 5.3 addresses this by extending the ballot to more than two candidates.

Another limitation of VVB is in its accessibility to blind and colour-blind voters. As it is designed, blind voters can vote with privacy and without assistance — unlike in many other voting systems. This is a limitation of the system, but the security of blind and colour-blind voters is still assured by the fact that their ballots are not identifiable. That said, if blind voters are present, care should be taken during the counting to check that the texture information corresponds to the written information.

We should also warn that as it is, VVB should not be used for elections with more than a few dozen voters. The ability to find one's visual secret has only been tested among 20 images, and not among a set of 200. It might be possible to generalise the method to handle more voters, but the viability of visual secrets in this context has not been tested.

5.3 Refinements on VVB and variants

A first type of modification would be to address the fact that the VVB shown here only allows binary elections. They are easily adaptable, however, by using circular ballots or polygonal ballots with $2n$ sides and n folding lines — one per candidate — inspired by what was proposed in [5].

There is also the question of the order of candidates on the ballot, which generally has a non-trivial impact[6]. Thanks to the symmetry and lack of favoured orientation when using abstract visual secrets, the vertical and horizontal lines are only identifiable by what is written on them. A quick improvement would then be to switch from "Vote 1/2" or "Vote A/B" to using "Vote 1" on one line and "Vote A" on the other, although it could confuse some voters[7]. A variant useful in certain cases would also be to have "Yes" and "No".

If needed, one could refine the ballot design to make them harder to describe (at the cost of potentially lowering the recognisability). For example, each ballot could have a thicker black border on the edges, to limit the possibility of catching part of someone else's picture if they bend the ballot. The visual secret could also be made circular (with a border covering the rest of the ballot), to make it harder to pinpoint what is in the corners (which was attempted by some subjects). Ultimately, a balance needs to be found between recognisability and describability, upon which it is hard to conjecture without further user studies.

Finally, there is the question of the scalability. The subjects in our experiment have shown their ability to remember multiple images. If we were to use 3 independent visual secrets on the same ballot, the probability of fraud detection

[6] The impact of candidate order has mostly been observed in large-scale elections where voters are not necessarily familiar with the candidates, and as such could have lower importance in boardroom settings[36].

[7] One small issue is that this is not directly compatible with the texture bumps as the symbols for 1 and A are identical in Braille, unless the complex numerical prefix is added.

could potentially reach 99.1% for each ballot (assuming all of the images are unique). An adversary coercing the subject into describing their visual secrets would have at best an 1.5% chance of correctly identifying all the images. This method requires more investigation, as a partial description could be enough to identify the voter depending on how the full set of visual secrets is constructed. Once again, we have balance problems that require more data to be resolved. That said, if a variation on the method could handle sets of 1 000 visual secrets, this kind of vote could happen on a much larger scale, with the verification happening by precinct, each set of visual secret being restricted to its precinct.

6 Concluding remarks

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

- the high recognisability of the pictures, with subjects having 80%+ chance of recognising their own secret ;
- 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 shows that cognitive responses can be used to design or improve low-tech voting protocols. We propose an example of protocol that could be used instead of traditional paper voting for boardroom voting. This protocol resists some of the central attacks against the standard boardroom voting practice, by preventing coercion and vote-selling (as people can't prove how they voted and ballots are harder to track than if they were handwritten) while discouraging the removal and replacement of ballots through the verification mechanism. Moreover, it can be used as a teaching tool to introduce people to verifiable voting at limited cost.

Visual secrets could be used as in our example, or potentially as a replacement for the identifying marks used in other verifiable voting systems such as sElect (which uses a 9 characters string chosen by the voter) [25]. Visual secrets could be a first tool allowing to tackle the numerous issues which motivated the designers of sElect not to address coercion. Another option would be to replace the identifying marks in protocols inspired by Ron Rivest's ThreeBallot [34, 4].

Outside of voting protocols, visual secrets could also be used within authentication mechanisms or online communication protocols. However, this is not trivial as the goals of visual secrets are quite different from picture passwords, their closest equivalent. In our case describability and short-term memorability are bigger concerns than speed or long-term memorability. We hope visual secrets will inspire other applications beyond boardroom voting.

6.1 Future work and open problems

The data in our experiment will be released publicly, and there are still a few leads that could be worth investigating (in an exploratory fashion) :

- Our initial attempts at making clustering analyses did not give us strong insights, but variations in memorability between the images within each series could be compared to a clustering of the descriptions.
- The memorability could be correlated with the time spent on the image description, and inversely correlated with the time spent on the description of other images (as they function as distractor tasks).
- The mobile interface used by 20% of subjects could have had an impact on performance, compared to a full-sized computer screen (potentially not because of the screen but because of how people interacted with the experiment).
- Describability could vary by subject, with some people being better at describing everything unambiguously, or it could mostly depend on the images assigned to the subjects.

Beyond this work on the dataset, this study raises multiple questions about refinements and extensions :

- Could the recognition process be fooled by images that are very similar, such as iterations on one basis made using generative adversarial networks [12] ?
- What would the performance become if the image series were composed of 100 images or more ?
- We did not measure confidence in the subject's choices when recognising, but it plays a role in the security aspect (in terms of false positives). What would be the recognition performance if we restricted to subjects who are sure of their choice ?
- 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 asked the subjects to describe from memory, a few minutes after viewing their pictures (which is closer to the real-life coercion scenario) ?
- How unambiguous would they 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 ? Could an adversary with some information on the image series create a teachable description method that would increase describability ?
- From a formal standpoint in both classical and quantum complexity, 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 it be possible to create tactile secrets (that could be also embossed in a ballot) ? Could

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?

Finally, our study of VVB is based on the performance of the original visual secret user study. A dedicated usability study on its performance in actual use could reveal new intuitions and leads for further improvements.

References

1. Arnaud, M., Cortier, V., Wiedling, C.: Analysis of an electronic boardroom voting system. In: Heather, J., Schneider, S., Teague, V. (eds.) E-Voting and Identify. pp. 109–126. Springer Berlin Heidelberg (2013)
2. Barrett, R.W.: Elephant in the boardroom: Counting the vote in corporate elections. Valparaiso University Law Review **44**, 125 (2009)
3. Bednarik, R., Kinnunen, T., Mihaila, A., Fränti, P.: Eye-movements as a biometric. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) Image Analysis. pp. 780–789. Springer Berlin Heidelberg, Berlin, Heidelberg (2005)
4. Blanchard, E., Selker, T.: Origami voting: a non-cryptographic approach to transparent ballot verification. In: 5th Workshop on Advances in Secure Electronic Voting (2020)
5. Blanchard, E., Selker, T., Sherman, A.T.: Boardroom voting: Practical verifiable voting with ballot privacy using low-tech cryptography in a single room (2019), https://hal.archives-ouvertes.fr/hal-02908421/
6. Blanchard, N., Malaingre, C., Selker, T.: Improving security and usability of passphrases with guided word choice. In: 34th Annual Computer Security Applications Conference, ACSAC. pp. 723–732 (2018)
7. Bojinov, H., Sanchez, D., Reber, P., Boneh, D., Lincoln, P.: Neuroscience meets cryptography: Designing crypto primitives secure against rubber hose attacks. In: 21st USENIX Security Symposium. pp. 129–141. USENIX, Bellevue, WA (2012)
8. Bradley, M.M., Lang, P.J.: Memory, emotion, and pupil diameter: Repetition of natural scenes. Psychophysiology **52**(9), 1186–1193 (2015)
9. Burton, C., Culnane, C., Schneider, S.: vvote: Verifiable electronic voting in practice. IEEE Security & Privacy **14**(4), 64–73 (2016)
10. Choudhury, B., Then, P., Issac, B., Raman, V., Haldar, M.: A survey on biometrics and cancelable biometrics systems. International Journal of Image and Graphics **18** (2018)
11. Cortier, V., Gaudry, P., Glondu, S.: Belenios: a simple private and verifiable electronic voting system. In: Foundations of Security, Protocols, and Equational Reasoning, pp. 214–238. Springer (2019)
12. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A.: Generative adversarial networks: An overview. IEEE Signal Processing Magazine **35**(1), 53–65 (2018)
13. Dufour, V., Coleman, M., Campbell, R., Petit, O., Pascalis, O.: On the species-specificity of face recognition in human adults. Cahiers de Psychologie Cognitive-Current Psychology of Cognition (2004)
14. Gebhardt Stenerud, I.S., Bull, C.: When reality comes knocking norwegian experiences with verifiable electronic voting. In: 5th International Conference on Electronic Voting 2012 (EVOTE2012) (2012)

15. Goldreich, O., Oren, Y.: Definitions and properties of zero-knowledge proof systems. Journal of Cryptology **7**(1), 1–32 (1994)
16. Groth, J.: Efficient maximal privacy in boardroom voting and anonymous broadcast. In: Juels, A. (ed.) Financial Cryptography. pp. 90–104. Springer (2004)
17. Haist, F., Shimamura, A.P., Squire, L.R.: On the relationship between recall and recognition memory. Journal of Experimental Psychology: Learning, Memory, and Cognition **18**(4), 691 (1992)
18. Jensen, W., Gavrila, S., Korolev, V., et al.: Picture password: A visual login technique for mobile devices. Tech. rep., National Institute of Standards and Technology (2003)
19. Johnson, M.H.: Subcortical face processing. Nature Reviews Neuroscience **6**(10), 766–774 (2005)
20. Jones, D.W.: A brief illustrated history of voting. University of Iowa Department of Computer Science. (2003), `http://homepage.divms.uiowa.edu/~jones/voting/pictures/`
21. Jonker, H.L., de Vink, E.P.: Formalising receipt-freeness. In: International Conference on Information Security. pp. 476–488. Springer (2006)
22. Kafkas, A., Montaldi, D.: 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), 1971–1989 (2011)
23. Krantz, J.H.: Psychological research on the net (2019), `https://psych.hanover.edu/research/exponnet.html`
24. Krantz, J.H., Reips, U.D.: The state of web-based research: A survey and call for inclusion in curricula. Behavior research methods **49**(5), 1621–1629 (2017)
25. Küsters, R., Müller, J., Scapin, E., Truderung, T.: select: A lightweight verifiable remote voting system. In: 2016 IEEE 29th Computer Security Foundations Symposium (CSF). pp. 341–354 (2016). https://doi.org/10.1109/CSF.2016.31
26. Li, S., Kot, A.C.: Attack using reconstructed fingerprint. In: IEEE International Workshop on Information Forensics and Security – WIFS. pp. 1–6. IEEE (2011)
27. McCorry, P., Shahandashti, S.F., Hao, F.: A smart contract for boardroom voting with maximum voter privacy. In: Kiayias, A. (ed.) Financial Cryptography and Data Security. pp. 357–375. Springer International Publishing, Cham (2017)
28. McCulley, S., Roussev, V.: Latent typing biometrics in online collaboration services. In: Proceedings of the 34th Annual Computer Security Applications Conference. pp. 66–76. ACSAC '18, ACM (2018)
29. Naber, M., Frässle, S., Rutishauser, U., Einhäuser, W.: Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. Journal of vision **13**(2), 11–11 (2013)
30. Nakajima, M., Yamaguchi, Y.: 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, Plzen-Bory, Czech Republic, February 4-8, 2002. pp. 303–310 (2002)
31. Naor, M., Shamir, A.: Visual cryptography. In: De Santis, A. (ed.) Advances in Cryptology — EUROCRYPT'94. pp. 1–12. Springer (1995)
32. Nelson, C.A.: The development and neural bases of face recognition. Infant and Child Development: An International Journal of Research and Practice **10**(1-2), 3–18 (2001)
33. NFB Jernigan Institute: The Braille literacy crisis in America. Tech. rep., National Federation of the Blind (2009)

34. Rivest, R.L., Smith, W.D.: Three voting protocols: Threeballot, vav, and twin. In: Proceedings of the USENIX Workshop on Accurate Electronic Voting Technology. pp. 16–16. EVT'07, USENIX Association, Berkeley, CA, USA (2007)
35. Shepard, R.N.: Recognition memory for words, sentences, and pictures. Journal of Verbal Learning and Verbal Behavior **6**, 156–163 (02 1967)
36. Sled, S.M.: Vertical proximity effects in the California recall election. Tech. rep., Caltech/MIT Voting Technology Project (2003)
37. Sluganovic, I., Roeschlin, M., Rasmussen, K.B., Martinovic, I.: Using reflexive eye movements for fast challenge-response authentication. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. pp. 1056–1067. CCS '16, ACM, New York, NY, USA (2016)
38. Smyth, J.D., Dillman, D.A., Christian, L.M., McBride, M.: Open-ended questions in web surveys: Can increasing the size of answer boxes and providing extra verbal instructions improve response quality? Public Opinion Quarterly **73**(2), 325–337 (2009)
39. Solvak, M.: Does vote verification work: Usage and impact of confidence building technology in internet voting. In: International Joint Conference on Electronic Voting. pp. 213–228. Springer (2020)
40. Wiese, H.: 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), 137 – 147 (2012)
41. Wolff, N., Kemter, K., Schweinberger, S.R., Wiese, H.: What drives social in-group biases in face recognition memory? erp evidence from the own-gender bias. Social Cognitive and Affective Neuroscience **9**(5), 580–590 (2014)