Received: 2 October 2018 | Revised: 6 November 2018 | Accepted: 13 November 2018

DOI: 10.1111/ejn.14278

TECHNICAL SPOTLIGHT

WILEY EJN European Journal of Neuroscience FENS

A novel tool for time-locking study plans to results

Matan Mazor[1,2] | Noam Mazor[3] | Roy Mukamel[1,4]

[1]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

[2]Institute of Neurology, University College London, London, UK

[3]Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel

[4]School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel

Correspondence
Matan Mazor, Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.
Email: matanmazor@outlook.com

Abstract

Often researchers wish to mark an objective line between study plans that were specified before data acquisition and decisions that were made following data exploration. Contrary to common perception, registering study plans to an online platform prior to data collection does not by itself provide such an objective distinction, even when the registration is time-stamped. Here, we adapt a method from the field of cryptography to allow encoding of study plans and predictions within random aspects of the data acquisition process. Doing so introduces a causal link between the preregistered content and objective attributes of the acquired data, such as the timing and location of brain activations. This guarantees that the preregistered plans and predictions are indeed specified prior to data collection. Our time-locking system does not depend on any external party and can be performed entirely in-lab. We provide code for easy implementation and a detailed example from the field of functional Magnetic Resonance Imaging (fMRI).

KEYWORDS

neuroimaging, preregistration, replicability, science

1 | INTRODUCTION

Preregistration of study plans and predictions prior to data collection sharpens the distinction between hypothesis-driven and exploratory phases of the scientific process (Munafò et al., 2017; Nosek et al., 2015). This effect takes place at two levels. First, at the single laboratory level, preregistering study plans encourages a healthy research workflow by providing an available reference for the original motivation, plans and predictions of the study. This desired effect does not depend on the preregistration mechanism, and to enjoy it,s one can write their study plans and predictions on a piece of paper and seal it in an envelope the evening before data collection is started. Secondly, at the scientific community level, study preregistration can be used as an objective marker for hypothesis based findings, as opposed to exploratory findings that can generate new hypotheses (De Groot, 2014). The sealed envelope scheme does not provide this objective marker, as from the reader's point of view there is no way to verify that the envelope was sealed before data collection, and not after it.

To enjoy the community-level benefits of preregistration, what is needed is an objective marker that preregistration was indeed performed before data collection. We refer to this as *time-locking*: proving that one event (preregistration) preceded a second event (data collection) in time. While online repositories such as osf.io, AsPredicted.org, and protocols.io can provide a time-stamp for the registration date, they cannot guarantee that data collection has not preceded this date. Importantly, only time-locking is of interest for preregistration. The date in which the registration has been submitted is of no interest without knowledge about the time of data acquisition relative to it (see Figure 1 for a comparison between

Abbreviations: EEG, electroencephalogram; fMRI, functional magnetic resonance imaging; OSF, open science framework; PRNG, pseudorandom number generator; RPR, reviewed preregistration; UPR, unreviewed prereg istration.

Edited by Guillaume Rousselet. Reviewed by Cyril Pernet and Antonio Schettino.

All peer review communications can be found with the online version of the article.

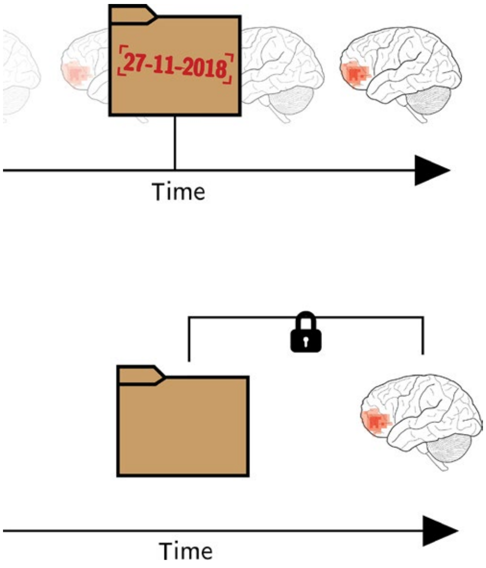

FIGURE 1 Upper panel: time-stamping. Platforms such as the open science framework can guarantee that the study protocol has been registered on a certain date, but they cannot guarantee that this date preceded data collection. Lower panel: time-locking. Our alternative scheme guarantees that regardless of when the preregistration was specified, data collection followed it in time

time-stamping and time-locking). In other words, reporting that study plans were preregistered using a review-free online platform does not offer anything beyond reporting that study plans were sealed in a closed envelope the evening of study commencement.

Here, we introduce a preregistration scheme that provides true time-locking while being performed in-lab, without the involvement of any third party. By using this scheme, researchers can provide objective and unequivocal proof that specific study plans and hypotheses have been specified before data acquisition, and therefore not after data exploration.

Our scheme is inspired by cryptographic protocols (Fiat & Shamir, 1986) and exploits random features in the experimental design to time-lock study plans. It is therefore applicable for experiments with aspects that can be determined in a pseudorandom fashion (such as the timing, order, or type of experimental events). Additionally, it is assumed that data is made available postpublication, either by uploading it to an online repository or by making it available upon request.

2 | THE PROPOSED METHOD

The purpose of our method is to make the claim "study plans and predictions were specified before data collection" a verifiable claim. To achieve this, we make the experimental design causally dependent on the specified study plans and predictions, such that every slight change to the registered plans results in a completely different experiment structure, and therefore in completely different patterns in the data. This makes the raw data a voucher for the preregistration validity.

To introduce this causal link, we make use of experimental randomization. Experimental randomization is most commonly obtained in a computerized way, using a pseudorandom number generator (PRNG) that is usually built in the programming language in use. A PRNG is a deterministic algorithm that generates a sequence of numbers whose properties approximate the properties of random numbers. Importantly, the output of a PRNG is not random at all, and in fact it is completely and deterministically determined by the PRNG initialization seed. This seed can be a number or a series of numbers, and is often chosen arbitrarily and shared in order to increase the reproducibility and transparency of the experimental results. Here, instead of an arbitrary choice of seed, we propose to choose the PRNG initialization seed to be strongly dependent on the study plans and predictions. By doing so, we make the experimental design causally dependent on the prespecified study plans and predictions.

Time-locking is thus performed by the researcher prior to data acquisition and can be verified by anyone at any later stage (e.g., during peer review or after publication). In the following section we outline the time-locking and verification stages, and then describe a real-life use of our method from the field of neuroimaging:

2.1 | Time-locking

1. Before data acquisition, a protocol file is saved to a *protocol folder* together with any available details which the authors wish to state in advance (such as number of planned measurements, predictions and analysis parameters that will be used). A script that uses a *pseudorandom number generator* (PRNG) to determine all random aspects of the experiment is also saved to the same folder. Importantly, only information that is included in the protocol folder will be regarded as preregistered and hypothesis-driven. Additional analyses that may appear in the manuscript will be considered exploratory.
2. A cryptographic hash function is applied to the protocol folder. This results in a sequence of bits that for all intents and purposes is unique to the protocol folder and its contents (the *protocol sum*). The protocol sum is guaranteed to be a number of a fixed length, regardless of the size of the protocol folder. This property will later allow us to use the protocol sum as the PRNG initialization seed. Furthermore, due to the special properties of cryptographic hash functions, it is infeasible to find a different protocol folder that will mapped to the same protocol sum, so the protocol-sum can be treated as a reliable fingerprint of the protocol folder.

3. The protocol sum is used as a seed to initialize the pseudorandom number generator (PRNG).
4. The PRNG is used to determine various random aspects of the experimental protocol, such as order and timing of experimental events, using the script that has been saved to the protocol folder in step 1. At this stage, the experimental randomization is completely determined by the contents of the protocol folder because the PRNG was initialized with the protocol sum.
5. As part of the publication process or beforehand, the protocol folder is uploaded to an online repository, and a link to this repository is included in the final manuscript. Raw experimental data is shared, either publicly or made available upon request.

In what follows, we will refer to this utilization of the PRNG for preregistration time-locking as the *pre-RNG scheme*.

In practice, steps 2–3 can be performed by replacing the call to the PRNG initialization function (Python's `random.seed(my_seed)`, Matlab's `rng(my_seed)` or R's `set.seed(my_seed)`) with our `pre-RNG` function (Python, R, and Matlab implementations are available to download from *github.com/matanmazor/prerng*). Unlike standard PRNG initialization commands that expect a number or an array of numbers as input, our pre-RNG function receives the path to the protocol folder as its argument. To use our scheme, the traditional call to the seed initialization function at the top of the experiment-generating scheme will be replaced with a line similar to `preRNG("D:/experiment/protocolFolder.zip")`.

The pre-RNG function calculates the protocol sum—a long number that for all intents and purposes is unique to this protocol folder (step 2). It then uses the protocol-sum to initialize the PRNG (step 3), and returns it as output for future reference. In cases where multiple randomization schemes are desired, the function can be called with an additional argument specifying the randomization serial number (see Discussion). In our implementations, we used the SHA-256 hash function that outputs a protocol-sum of 256 bits for any arbitrary length input (NIST, 2002). Crucially, it is infeasible to find two inputs that are mapped to the same output by SHA-256.

2.2 | Verification

The pre-RNG time-locking introduced a causal link between the acquired data and the content of protocol folder via a chain of dependencies.

1. The dependency of the acquired data on random components of the experimental design (red arrow 1 in Figure 2) is a prerequisite for the use of this scheme. It is assumed that different experimental designs will yield different patterns in the data, and that post hoc manipulation of the data with the purpose of making it compatible with an alternative protocol is detectable. These conditions are met in most experimental designs that measure a continuous variable, for example, studies involving neuroimaging methods, such as EEG and fMRI.
2. The dependency of random components of the experimental design on the protocol sum (red arrow 2 in Figure 2) was obtained through the initialization of the PRNG with the protocol sum as seed. This dependency is tight, since the behavior of the PRNG is deterministically set given a particular seed, and different seeds result in different behaviors (Matsumoto, Nishimura, Hagita, & Saito, 2005).
3. The dependency of the protocol sum on the protocol folder (red arrow 3 in Figure 2) was obtained through the use of

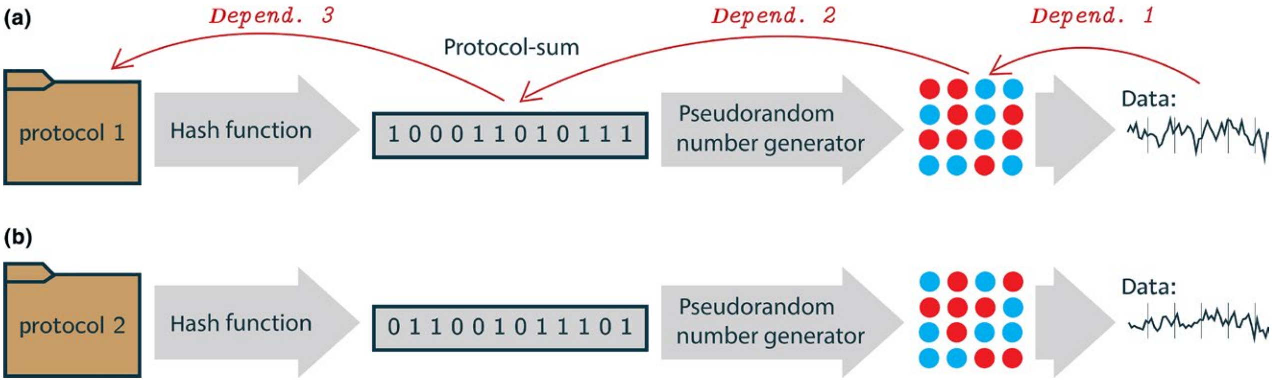

FIGURE 2 The pre-RNG scheme. Gray arrows (left to right) represent the temporal order of events. Red arrows (right to left) represent the causal structure. Registration time-locking is obtained by making the acquired data (represented as line plots) dependent on the protocol folder via specific random components of the experimental design (represented as blue and red dots). Panels (a) and (b) represent two alternative protocol-folders (for example, including alternative predictions). The slightest difference in the content of the two folders results in completely different randomizations and therefore different structures of data variability. This chain of dependencies time-locks the preregistration with respect to data acquisition

a cryptographic hash function. Such functions map arbitrary length inputs to sequences of bits of a fixed length such that it is infeasible to find two inputs, in our case, two protocol folders, that are mapped to the same sequence of bits.

Altogether, this chain of dependencies enforces a causal link and thus time-locks the acquired data with respect to the contents of the protocol folder, be it predictions, analyses plans or parameter specifications. Due to this causal link, raw experimental data should be in line with the randomness incurred by the PRNG that has been initialized with the protocol-sum as seed. This can be verified by analyzing the shared data according to the analysis plans specified in the protocol folder, by visual inspection, or using any other data-based verification tests chosen by the verifier — be it an editor, a reviewer, or an interested reader.

Using our scheme provides an objective marker for the integrity of the report. But even more importantly, by using this scheme authors communicate to their readers that it would be irrational of them to be dishonest about their original plans and predictions. To change the contents of the protocol folder retrospectively, authors will have to either manipulate their data, to lie about what actually happened in the experiment, or to lie about what their experiment randomization code is doing (see immunity to hacking strategies in the discussion). All of these lies can potentially be detected by the community and cause much more inconvenience than simply admitting that certain decisions were made post hoc, as is often the case in most research projects.

3 | EXAMPLE

To demonstrate the use of the pre-RNG scheme, we describe a scenario involving a researcher (Alice) and an interested reader (Bob). Alice examined the involvement of the cerebellum in voluntary hand movement and committed to her study plans using our pre-RNG scheme. Bob wants to verify that certain findings that are especially relevant to his own research are indeed hypothesis-driven, as reported. Alice's experiment was time-locked and physically conducted in our laboratory for demonstration. Alice's paper and Bob's verification are both included in the Appendix S1.

Bob downloads the study protocol folder from the link provided in the manuscript and in it he finds a methods section specifying Alice's choice to restrict her analysis to the cerebellum. He now wants to verify that this protocol folder has really been specified prior to data collection, rather than post hoc. To do so, Bob runs the pre-RNG function on the protocol folder. As Bob and Alice both applied the same deterministic function to the same input, Bob obtained a protocol-sum that is identical to the one obtained by Alice (dependency number 3). The pre-RNG function automatically initialized the PRNG on Bob's computer with this protocol sum, such that Bob's computer will now generate the same sequence of pseudorandom numbers as did Alice's computer when designing her experiment. Bob then uses the Python script that he found in the protocol folder to generate a pseudorandom sequence of experimental events. As Bob and Alice obtained an identical protocol-sum and since for a given seed PRNGs are deterministic, Bob obtains the same sequence of events that was used by Alice in the actual experiment (dependency number 2). Given the high number of possible orders of events in Alice's experiment, Bob concludes that the likelihood of obtaining a particular sequence of events by chance is very small ($<10^{-20}$), and therefore that the probability that a different PRNG seed would have resulted in a similar order of events is negligible.

In order to verify that the data reflects this randomization (dependency number 1), Bob writes to Alice and asks for the raw experimental data. The evoked responses in Alice's raw data should correspond to the pseudorandom order of experimental events that Bob obtained by initializing the PRNG with Alice's protocol folder and running her randomization script. To test this, Bob decides to perform a contrast between right and left hand movements, assuming this specific order of events. Note that Bob is free to choose whatever verification analysis he finds fit and is not limited to the analysis reported by Alice (for additional verification steps he can use, see Appendix S1). The resulting map is in line with Bob's prior knowledge of robust lateralized brain activations in primary sensorimotor cortices (see Figure 3). As fMRI data are highly affected by the specific order of events, the alignment of the acquired data with the pseudorandom order of events is a reliable voucher for the preregistration validity. Bob is now convinced that the randomization induced by the protocol-sum is in line with the data, and therefore that Alice's specification of analysis plans and predictions in the protocol folder was genuinely made prior to data acquisition.

4 | DISCUSSION

In the above example, the pre-RNG scheme allowed Alice to provide objective support for her claim that certain choices have been made *prior* to data collection, without sharing her study plans with any external party at an early stage. This would not have been possible in any other preregistration implementation.

Unreviewed preregistration platforms (UPR; van't Veer & Giner-Sorolla, 2016) such as AsPredicted and the OSF serve as an open vault for researchers to submit their study plans, and thus increase transparency and openness, and push toward the adoption of healthier research workflows. These platforms can also provide time-stamping of the

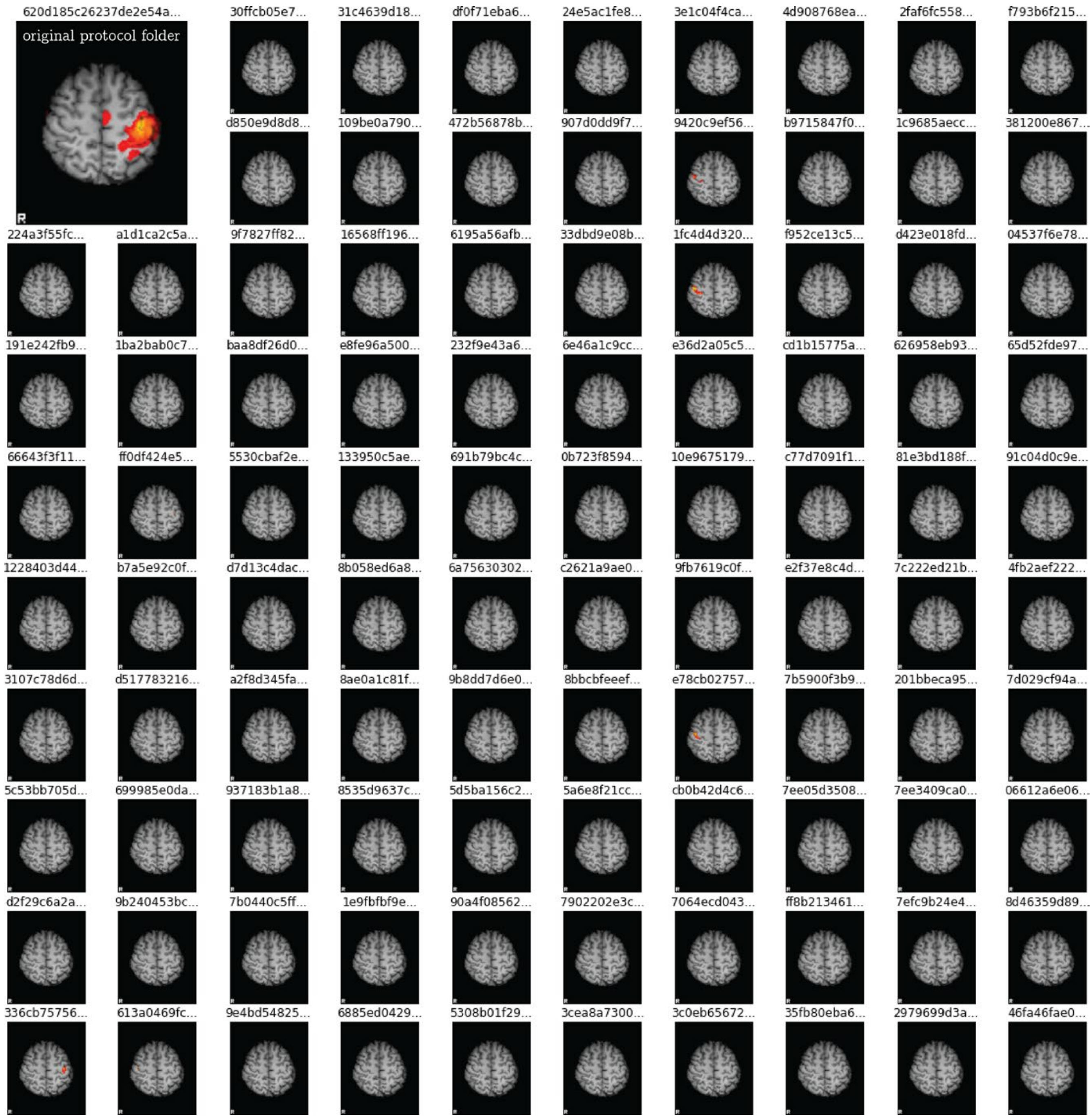

FIGURE 3 A contrast between right and left hand movements for Alice's data. Top left corner: Here Bob performed the contrast assuming the order of events as determined after applying the pre-RNG function to the protocol folder. All other maps were generated by applying the pre-RNG function to slightly altered protocol folders (by adding hash symbols to the end of one file in the folder). This slight change in protocol folder resulted in activation maps weaker by orders of magnitude compared to the one obtained from the original protocol folder. The first bits of the protocol sums are presented above each map

registration process, but they cannot guarantee that the registration of study protocol is indeed time-locked to precede data acquisition.

To date, time locking of protocol registration can only be obtained by introducing an additional peer-review step at an early stage of work (reviewed preregistration; RPR; Chambers, 2013). Reviewed preregistration time-locking relies on the premise that authors will be less likely to preregister studies for which data has already been collected when knowing that reviewers might request changes in the experimental design. Some RPR schemes have the advantage of facilitating the publication of null results by committing to publish regardless of outcome — a feature that is not supported by our in-laboratory approach. Nonetheless not

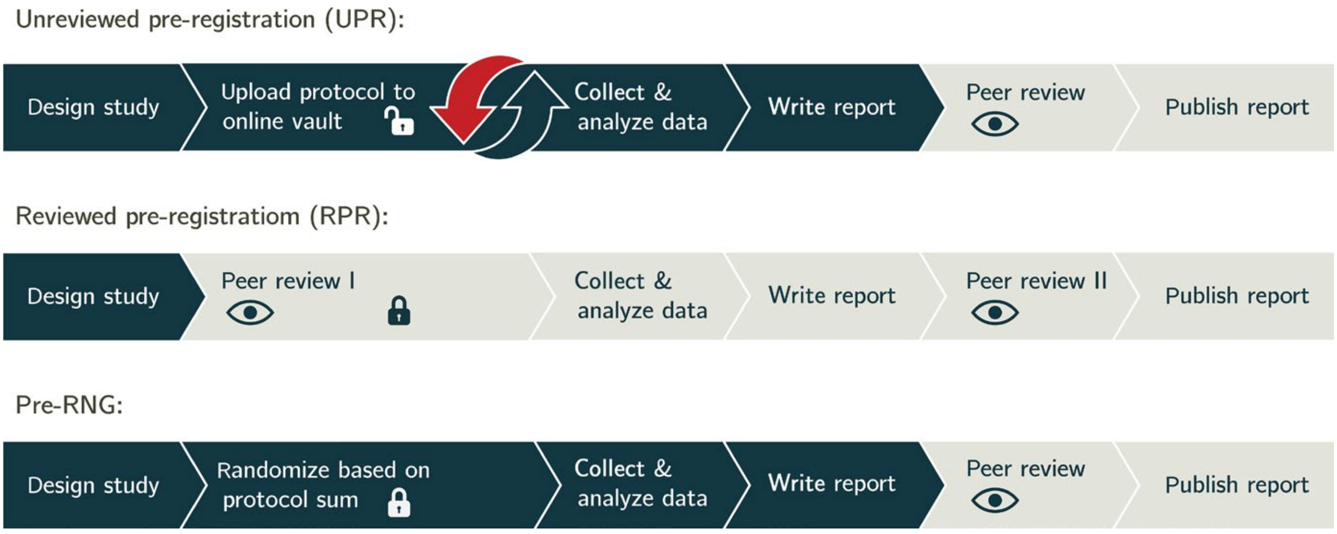

FIGURE 4 The three preregistration schemes. The transition from dark blue to light gray indicates the first mandatory exposure of the research protocol to a third party. Lock icons represent the commitment to a specific research protocol. The red arrow in the UPR scheme represents the loophole allowing one to "pre-register" research plans even after data collection and exploration. Our pre-RNG scheme is time-locked but does not require early exposure

all research projects are suitable for early-review schemes, which necessarily require early exposure of study plans to external reviewers.

Here, we propose a preregistration scheme that provides time-locking without the involvement of journals or peer reviewers early on in the process. Our approach requires authors to be willing to share their data upon publication—a practice that is becoming more and more prevalent (and sometimes even a requirement by funding agencies and journals) in behavioral and neuroimaging studies. It also requires that some aspects of the experimental design be determined pseudorandomly, which is also very common in psychological and neuroscientific research.

Our scheme can be easily incorporated into UPR online platforms such as OSF, such that upon material registration a protocol-sum is generated and given to the researcher, to be used as a seed for experimental randomization. Similarly, RPR schemes such as Registered Reports can ensure time-locking by providing researchers with a protocol-sum after completing the first review phase. This way, time-locking is guaranteed regardless of whether reviewers ask changes to be made to the study design or not.

In what follows we describe a few technical but important aspects and properties of the pre-RNG scheme, and outline how it can be applied to a variety of cases. We start by describing how our scheme can be used with studies that involve multiple, insufficient or constrained randomizations, we then move to describe how one may go about verifying a pre-RNG registration, and finally demonstrate the immunity of our scheme to different hacking strategies.

4.1 | Studies involving multiple randomizations

In cases where multiple randomizations are needed for different repetitions of the same experiment, such as in the case of multiple subjects, the pre-RNG function can be called with an optional argument specifying the subject's serial number. The serial number will be appended to the protocol sum, and the SHA256 function will be applied to the resulting string. The new sum will be used to initialize the PRNG. This guarantees that (a) different repetitions will be initialized with completely different seeds and that (b) for all subjects, the randomization is fully dependent on the protocol folder. This option is supported by the accompanying implementations.

4.2 | Studies with insufficient randomization

Some experimental designs do not include any random component, and in others the entropy of the experimental randomization is not sufficient to effectively time-lock the protocol folder (for example, randomizing the order of two experimental blocks: A, B or B, A). To use the pre-RNG scheme in such cases, one can add an additional randomized experimental phase only for time-locking. For example, neuroimaging experiments can begin with a short block of events that give rise to robust sensory or motor activations, in random order and timing.

4.3 | Constrained randomization

Pre-RNG can be used even in the presence of constraints to the randomization scheme, as long as the probability of each specific pseudorandom design is very small. For example, assuming that the probability of selecting a specific design at random out of the constrained pool of designs is at most $\frac{1}{1{,}000}$, experiment with 12 subjects and 4 independent runs per subject will get selected randomly with a probability of $\frac{1}{1{,}000^{4x12}}$. In other words, it will be more likely to randomly pick the same atom twice from all atoms in the universe, than to pick twice the same experimental design.

We tested this assumption for the case of efficiency-optimized functional MRI event-related designs. We used the popular optseq2 (Dale, 1999) to generate design matrices for an event-related functional MRI design, following the second example in Harvard's online tutorial (http://surfer.nmr.mgh.harvard.edu/optseq/optseq-practical.txt). We looped over the same command 350,000 times when the only thing that changed was the PRNG initiation seed, and haven't found one collision (i.e. two randomization seeds that are mapped to the same design). Thus, with very high confidence, there is no design matrix that is selected with a probability of $\frac{1}{1{,}000}$ or more.

4.4 | Verification as comparing alternative randomization schemes

To verify the dependence of the data on the PRNG initialization, the verifier can generate a null distribution of results assuming different randomizations that were obtained using arbitrary seeds, while keeping the data constant. For example, the verifier can derive the result of a contrast between two experimental conditions assuming the random order of events that is dictated by the original protocol folder, and compare it to the result of the same contrast when assuming other possible orders generated by initializing the PRNG with alternative seeds. While this is practically similar to the use of permutation testing for nonparametric inference, it is conceptually different: here the effect of interest is assumed to be known, and inference is made on the true experimental randomization that was used to generate the data.

Evidence that the true protocol folder gives rise to randomization that is sufficiently similar to the one that maximizes the effect of interest (or the likelihood of the data) can be used to corroborate the validity of the registration process. In other words, the credibility of the preregistration is made quantifiable.

4.5 | Immunity to hacking strategies

4.5.1 | Changing the analysis protocol in retrospect

Any changes made to the protocol folder after data acquisition will change the randomization scheme altogether: the protocol-sum will be completely different, and the pseudorandom experiment structure will change accordingly. This will necessarily break the alignment of the data with the experimental protocol. Such misalignment will be easily detected by readers who wish to validate the pre-registration.

4.5.2 | Registering multiple study protocols for the same data set

As all random components in the experiment are determined by the same PRNG and using the same seed, only one protocol folder can be associated with a particular study. This folder can of course include more than one possible protocol (e.g., more than one analysis schemes), but this will demonstrate that the researchers did not commit to a single protocol before data acquisition, making the analyses effectively post hoc.

4.5.3 | Reporting only successful repetitions

To date, no preregistration scheme is immune to selective reporting of subjects or experimental repetitions. In theory, authors can repeat the same experiment over and over until observing the desired effect, and report only the last, successful, repetition. Similarly, in the case of multi-subject experiments, authors can run as many subjects as they wish, only to report a subset of these subjects whose data aligns with the prior hypothesis. Our scheme is not immune to such misconduct. However, this concern can be alleviated by introducing a dependency between consecutive subjects. One way to introduce such a dependency is to call the pre-RNG function with the experimental data of subject *n-1* when determining random aspects of the experimental design of subject *n*. This way, the data acquired from subject *n* is dependent upon the data acquired from all the previous subjects and the original protocol folder, making it impossible to report a subset of the subjects without breaking the alignment between the data and the experimental randomization.

5 | SUMMARY

By providing a registration time-locking method that does not involve early exposure of study plans, we hope to encourage research laboratories to preregister predetermined aspects of their studies and by doing so delineate a clearer border between hypothesis-driven and exploratory findings in their final report. This is an important step in making the scientific process more open and transparent, and increasing replicability of scientific findings.

ACKNOWLEDGMENTS

The authors thank Halely Balaban, Roni Maimon, Tal Golan, and Bat-Sheva Mazor for their helpful comments.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

AUTHOR CONTRIBUTIONS

N. Mazor came up with the original idea. N. Mazor and M. Mazor implemented the pre-RNG method in code. M. Mazor designed and conducted the fMRI experiment and analyzed the results under the supervision R. Mukamel, M. Mazor, N. Mazor and R. Mukamel wrote the manuscript. All authors approved the final version of the manuscript for submission.

ORCID

Matan Mazor https://orcid.org/0000-0002-3601-0644
Roy Mukamel https://orcid.org/0000-0001-9359-8950

REFERENCES

Chambers, C. D. (2013). Registered reports: a new publishing initiative at Cortex. *Cortex*, *49*(3), 609–610. https://doi.org/10.1016/j.cortex.2012.12.016

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human brain mapping*, *8*(2–3), 109–114. https://doi.org/10.1002/(ISSN)1097-0193

De Groot, A. D. (2014). The meaning of "significance" for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han LJ van der Maas]. *Acta psychologica*, *148*, 188–194. https://doi.org/10.1016/j.actpsy.2014.02.001

Fiat, A., & Shamir, A. (1986). How to prove yourself: Practical solutions to identification and signature problems. *Advances in Cryptology—CRYPTO'86* (pp. 186–194). Berlin, Heidelberg: Springer.

Matsumoto, M., Nishimura, T., Hagita, M., & Saito, M. (2005). Cryptographic mersenne twister and fubuki stream/block cipher. *IACR Cryptology ePrint Archive*, 165.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., & du Sert, N. P., … Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

National Institute of Standards and Technology (NIST) (2002). *FIPS 180-2*. Retrieved from https://csrc.nist.gov/csrc/media/publications/fips/180/2/archive/2002-08-01/documents/fips180-2.pdf

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Contestabile, M. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

van't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology*, *67*, 2–12. https://doi.org/10.1016/j.jesp.2016.03.004

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Mazor M, Mazor N, Mukamel R. A novel tool for time-locking study plans to results. *Eur J Neurosci*. 2018;00:1–8. https://doi.org/10.1111/ejn.14278