Cimiano P., Pietsch C., & Wiljes C. (Eds.) (2021). Studies in Analytical Reproducibility: the Conquaire Project. Bielefeld. https://doi.org/10.4119/unibi/2942780

9 | Reproducing the analysis of an experiment in sequential visual processing

Rebecca Foerster[2], Philipp Cimiano[1], Werner X. Schneider[2]

1 – Faculty of Technology & Cognitive Interaction Technology Excellence Center (CITEC), Bielefeld University

2 – Faculty of Psychology and Sports Science, Department of Psychology & Cognitive Interaction Technology Excellence Center (CITEC), Bielefeld University

Abstract

This chapter describes a case study in reproducing work conducted by the neuro-cognitive psychology research group at Bielefeld University in the area of sequential visual processing. In particular, we describe our effort to independently reproduce the results obtained via the experiment conducted in the paper *'Expectation violations in sensorimotor sequences: shifting from LTM-based attentional selection to visual search'* [1]. The research of the group focuses on the area of visual attention, eye movements, working memory, transsaccadic learning, and sensorimotor learning. The group works on understanding visual processing in humans via controlled behavioral experiments in laboratory environments alongside real-world studies. The main result of the article mentioned above was the finding that expectation violations in a well-learned sensorimotor sequence in humans caused a regression from LTM-based attentional selection to visual search. We describe in this paper our efforts to independently reproduce these results. We conclude that this case is a case of limited analytical reproducibility in that results are reproducible by relying on SPSS as in the original data analysis or by adapting analysis codes to open-source software packages such as R. The data and scripts for this project are available at `https://gitlab.ub.uni-bielefeld.de/conquaire/neurocognitive_psychology`.

Keywords

attention, eye movements, long-term memory (LTM), visual search, sensorimotor action, expectation discrepancy

9.1 Introduction

The neuro-cognitive psychology group at Bielefeld University is mainly concerned with research on visual attention, visual working memory, eye movements, transsaccadic learning, and sensorimotor control and learning. A first key issue is to understand how humans employ their visual attention to control movements. A second issue refers to elementary neuro-cognitive mechanisms as well as to group differences between healthy individuals and patients in visual attention and working memory. In order to achieve these goals, the neuro-cognitive psychology group conducts controlled behavioral experiments in the laboratory as well as real-world studies. The experiments often afford highly precise presentation durations of visual material, which are achieved by employing CRT screens, G-sync LCD monitors [2], high-speed projectors or head-mounted virtual reality devices [3]. Behavioral responses (e.g., letter reports, key presses), eye movements (static and mobile eye tracking), and hand movements (motion tracking, mouse cursor tracking) as well as video data and EEG data are recorded.

Within the Conquaire project, the publication by Foerster and Schneider entitled *'Expectation violations in sensorimotor sequences: Shifting from LTM-based attentional selection to visual search'* [1] was chosen to be reproduced. In that article, the consequences of violating long-term memory (LTM) based expectations about a learned sensorimotor sequence were investigated. Especially for well-practiced sequential sensorimotor actions, such as *driving, making a sandwich or performing sports*, LTM expectations have an important role because they guide the necessary task-adapted sequence of covert shifts of attention, eye movements, and hand and body movements [4, 5, 6, 7, 8]. In the study reported in the manuscript, it was investigated which consequences arise for eye and hand movement control when a learned visuospatial configuration (fixed sequence of spatially distributed mouse clicks) was unexpectedly changed.

Results revealed that the changes of action-irrelevant visual features of the configuration had no effect, neither on hand nor eye movements. In contrast, changes of the visuospatial configuration that forced participants to update their learned sensorimotor sequence partly affected both hand and eye movements. Such changes slowed down the demanded action, they elicited visual search-like scanning that replaced the previously LTM-controlled eye movements, and they reduced the eyehand synchrony. These effects were neither limited to the changed stimuli nor to actions on them.

We describe the specific experimental settings of the original work in Section 9.2. After this description, we provide details on how we attempted to reproduce the main results of the above mentioned paper.

9.2 Methods

Here, we describe the experimental setting and the methods used in the experiments conduced in the paper *'Expectation violations in sensorimotor sequences: shifting from LTM-based attentional selection to visual search'* [1].

9.2.1 Experiment settings and Data acquisition pipeline

In order to investigate the effects of action-relevant and action-irrelevant expectation violations on eye and hand movements in [1], the following experimental design was adopted. Forty right-handed participants (mean age of 25 years, 14 male, 26 female) were recruited at Bielefeld University, with normal or corrected-to-normal vision, to participate in the computer experiment.

All participants were first trained for 60 trials to click as fast as possible with a computer mouse in ascending order on eight numbered unique shapes on a computer screen (1-8). Importantly, the spatial configuration of the numbered shapes was constant over the course of the 60 trials (Figure 9.1), so that participants could learn and automatize the visuospatial configuration of the numbered shapes as well as the clicking sequence.

Typically, an eye movement to a location preceded each clicking action on a location. Thus, participants adopted LTM expectations about the visuospatial characteristics and an LTM-based control of visual attention and eye movements. After the 60th trial, we violated these visuospatial expectations unannounced, so that the 20 consecutive trials had a different configuration.

The 40 participants were divided into four experimental groups of 10 participants each, depending on the changed features during the 20 change trials.

- In the **shape-change group**, the shapes (circle and a plus sign) surrounding the numbers *3* and *6* switched position.
- In the **number-change group**, the numbers *3* and *6* changed position without changes in the surrounding shapes.
- In the **object-change group**, the numbers *3* and *6* switched position together with their surrounding shapes, so that the objects remained constant, e.g., a *plus 3* and a *circle 6*.
- In the **no change control group**, no switch was introduced.

As the **shape change** does not require a change of the learned clicking action, we call this an **action-irrelevant change**. As the **number and object changes** do affect the learned clicking sequence, we call these changes **action-relevant**.

In order to investigate how previously learned expectations and sensorimotor sequences can be re-initiated, 20 reversion trials followed the 20 change trials, in which the configuration was the same as during the 60 pre-change trials for each participant.

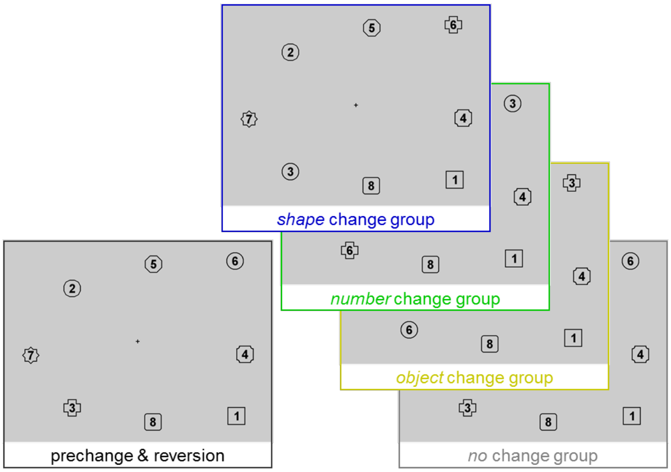

Figure 9.1: Computer display in the clicking task experiment

Figure 9.1 shows the display during the clicking task in the prechange (left), change (right), and reversion (left) phase of the experiment for the even participants of the four change groups (shape, number, object, no). Odd participants started with the plus three in the upper right position and the circle six in the lower left position.

Throughout the whole experiment, cursor movements on the CRT computer screen (ViewSonic Graphics Series G90fB, 19 inch color monitor @ 1024 x 768 pixels) were recorded with 100 Hz and participants' right gaze positions were recorded with an Eyelink 1000 desktop-mounted eye-tracker (SR Research, Ontario, Canada) with 1000 Hz. A standard computer mouse and an extra-large mouse-pad (32 x 88 cm) were used. A forehead and chin rest was used to fix participants' viewing distance at 71 cm. All stimuli were presented in black on a grey background. The cursor was a black dot subtending approximately 0.45° v.a. (degrees of visual angle) in central vision. A black plus sign with a height and width of 0.45° v.a. was presented in the screen center. The numbers were presented in bold Arial font with a font size of 35. Each number was surrounded by one unique shape with a diameter of about 2.18° v.a. in central viewing. The pre-change arrangement of the numbered shapes was generated randomly with the prerequisite that each outer field of an imagined 3 x 3 grid contained one shape and that the distance between shapes as well as the distance to the screen border was at least 2.18° v.a (border to border). For the generated configuration, the actual minimal distance happened to be 7.20° v.a. between the shapes containing numbers 1 and 4.

All participants saw numbers 1, 2, 4, 5, 7, and 8 in the same individual shapes and at the same location (Figure 1). Even participants saw a plus 3 in the lower-left corner and a circle 6 in the upper-right corner during the pre-change phase, while odd participants began with the switched position of plus 3 and circle 6. Each experimental group consisted of an equal number of odd and even participants, so that possible variations in the difficulties of the trajectories were cross-balanced.

The experiment was controlled by SR Research's Experiment Builder software. A nine-point eye-tracking calibration and validation procedure with an averaged accuracy criterion of 1.0° v.a. preceded the experiment. Calibration accuracy was checked before each trial on the basis of a central fixation on a black ring (0.48° v.a. outer size, 0.12° v.a inner size). Calibration was repeated if necessary.

After reading an initial written instruction on the computer screen, participants completed an example pre-change trial before the experiment started. This practice trial was not included in the analysis. Clicks were counted as correct within a diameter of 3.27° v.a. around a target's center. An incorrect click was followed by a low-pitched tone. After all eight objects were clicked sequentially in the correct order, participants were informed about their trial-completion time via a feedback display. After every block of 10 trials, a display informed participants about the number of blocks completed out of the total number of blocks. Participants started a block and a trial by pressing the space bar. All participants completed the experiment within 40 minutes.

Fixation detection

Fixations were detected by SR Research's default velocity algorithm (not a blink, velocity <30° v.a./s and acceleration $<$ 8000° v.a./s2). The following dependent variables were analyzed:

- trial-completion time,
- number and size of errors,
- number and duration of fixations,
- scan-path length,
- cursor-path length, and
- eyecursor distance.

Error size was measured as the Euclidean distance (° v.a.) from the center of the actual target to the incorrectly clicked location. Scan-path and cursor-path lengths were calculated as 100-Hz cumulative inter-sample distances. Eyecursor distances were calculated as 100-Hz intra-sample distances. For pre-change analyses, repeated measures analyses of variances (ANOVAs) with the within-subject factor block (6) were calculated for each dependent variable over all groups.

9.2.2 Methods applied to analyze the experiment data

For the change analyses, mixed design ANOVAs were calculated with change group (shape, number, object, no) as *between-subject factor* and phase (pre-change, change, reversion) as *within-subject factor*. For more fine-grained analyses, further ANOVAs were calculated including sub-action (8), location (8), and fixation type (searching, guiding, checking) as within-subject factors. *Guiding fixations* are fixations on current action goals, also known as *sequence or directing fixations* [9, 10, 11, 1, 8]. In the study, guiding fixations were operationalized as fixations to the numbered shape that was the current clicking target. *Checking fixations* are fixations to objects and locations that have already been acted on in the nearer past [10, 11, 1, 8]. In the study, checking fixations were operationalized as fixations to numbered shapes that had already been clicked correctly. *Searching fixations* are fixations to objects and locations that are currently not action-relevant, were not relevant shortly before, but might become relevant in the later future [9, 11, 1]. In the study, searching fixations were operationalized as fixations to numbered shapes that had not yet been clicking targets. Fixations were counted as falling on a numbered shape within an area of 3.27° v.a. around its center.

A LTM mode of visual attention is characterized by about one guiding fixation per sub-action of the task and nearly no checking or searching fixations, while searching fixations are indicative for visual search. Paired t-tests were conducted in case of significant two-way ANOVA interactions to reveal whether the values of two phases were significantly different across groups, sub-actions or locations. Violations of sphericity were corrected using Greenhouse-Geisser ϵ, but uncorrected degrees of freedom were reported to facilitate reading. A chance level of 0.05 was applied. Data preprocessing was conducted with MATLAB 2012a, data aggregation and diagrams were compiled in Microsoft Excel 2010, and statistical analyses were conducted with IBM SPSS Statistics 22.

The shape change did not affect any dependent variable significantly, neither when comparing the shape-change to the control group nor when comparing the shape-change phase values to the pre-change values. However, all dependent variables were strongly affected in the number and object change group with their values during the change phase differing from the pre-change values as well as from the control group. Specifically, participants of the number and object change group were slower, made more fixations, had longer scan-paths and cursor-paths and a larger eye-cursor distance during the first change trial than during the last pre-change block (pre-change baseline) as well as compared to the participants in the control group. Note that other pre-change baselines did not change the result pattern. Statistics can be viewed in the original paper.

Moreover, the type and size of the effects did not differ significantly between the number change group and the object change group. Therefore, these two groups were aggregated to one action-relevant change group for further analyses. The main results of these analyses were concerned with the number of fixations and fixation types performed by the action-relevant change group during the change compared to the pre-change phase.

To reveal which mode of attentional selection was predominantly applied before and after the action-relevant change, a repeated measures ANOVA was computed for the number of fixations with phase (pre-change, change) and fixation type (checking, guiding, searching) as within-subject factors. The analysis revealed significant main effects and a significant interaction on the number of fixations (phase: $F(1{,}19) = 23.97$, $p < 0.001$, $\eta_p^2 = 0.56$; type: $F(2{,}38) = 89.23$, $p < 0.001$, $\eta_p^2 = 0.82$; interaction: $F(2{,}38) = 21.49$, $\epsilon = 0.77$, $p < 0.001$, $\eta_p^2 = 0.53$; Figure 9.2, top). Paired t-tests revealed that the interaction was due to the fact that the change increased the number of searching fixations ($t(19) = 7.31$, $p < 0.001$), while there was no significant effect on the number of checking ($t(19) = 1.81$, $p = 0.09$) or guiding ($t(19) = 0.29$, $p = 0.80$) fixations. With nearly no checking (0.26) or searching (0.91) fixations per pre-change trial, LTM-based attention seemed to be the dominant mode of attentional selection after having learned the clicking sequence on the constant visuospatial configuration. The increase to about 5 searching fixations in the trial with the action-relevant number switch indicates a re-initiation of visual search.

Given that number 3 was no longer in the expected location, searching for the 3 when having to act on it is inevitable. Therefore, the question arises, whether searching is restricted to this action 3 or whether visual search is also initiated for other actions. To reveal whether searching fixations were differently prominent for different sub-actions, a repeated measures ANOVA was conducted for the number of searching fixations with the within-subject factors phase (pre-change, change) and action (1-7). The analysis for the number of searching fixations (Figure 2, middle) revealed significant main effects of phase ($F(1{,}19) = 53.43$, $p < 0.001$, $\eta_p^2 = 0.74$) and action ($F(6{,}114) = 20.85$, $\epsilon = 0.42$, $p < 0.001$, $\eta_p^2 = 0.52$) as well as a significant interaction ($F(6{,}114) = 19.74$, $\epsilon = 0.48$, $p < 0.001$, $\eta_p^2 = 0.51$). Paired t-tests revealed that the number of searching fixations was significantly increased during actions 3 ($p < 0.001$) and 4 ($p < 0.01$). Thus, searching fixations increased as soon as the first location-shifted number became the action target, but their increase was not limited to this action.

Do participants really search or is the increase in searching fixations completely explained by the fixations to the old position of number 3, i.e. the new number 6, which is by definition a not yet completed target? A repeated measures ANOVA for searching fixations with the within-subject factors location (2-8) and phase (pre-change, change) revealed two main effects and a significant interaction (location: $F(6{,}114) = 7.32$, $\epsilon = 0.28$, $p < 0.001$, $\eta_p 2 = 0.28$; phase: $F(1{,}19) = 44.80$, $p < 0.001$, $\eta_p^2 = 0.70$; interaction: $F(6{,}114) = 5.80$, $\epsilon = 0.45$, $p < 0.01$, $\eta_p^2 = 0.23$; Figure 2, bottom). Paired t-tests revealed that significantly

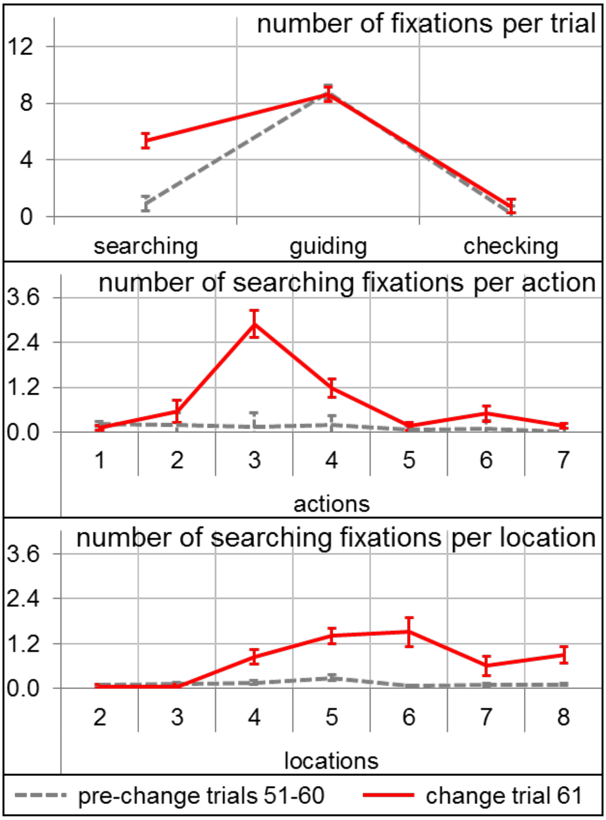

Figure 9.2: **Top panel**: Number of fixations per trial of the three fixation types searching, guiding, and checking. This is Fig.4 (top) from the original paper. **Middle panel**: Number of searching fixations per action (1-7). No searching fixations can be made during action 8 as there are no future targets. This panel is Fig. 5c from the original paper. **Bottom panel**: Number of searching fixations per location (2-8). No searching fixations can be made on location 1, as this location is never a future target. This figure is not in the original paper!

more searching fixations went to the locations 4-6 and 8 (all ps $< .0.5$), but not to the locations 2 ($p = 0.48$), 3 ($p = 0.24$), and 7 ($p = 0.06$). Thus, the increase in searching fixations is not limited to the new location of the 6, indicating that participants really search through the display.

To reveal how long it takes to incorporate the new clicking sequence, the number of searching fixations during the subsequent change trials was compared via paired t-tests to the pre-change baseline. Results revealed significantly more searching fixations compared to the pre-change baseline in the first 15 repetitions of the changed number display (all ps < 0.05; Figure 9.3). This result indicates that far more than a single trial is necessary to update parts of a learned sensorimotor sequence. Thus, sensorimotor updating of unexpected target locations can clearly be differentiated from surprise effects to unexpected visual items as surprise effects are typically very short-lived [12, 13, 14].

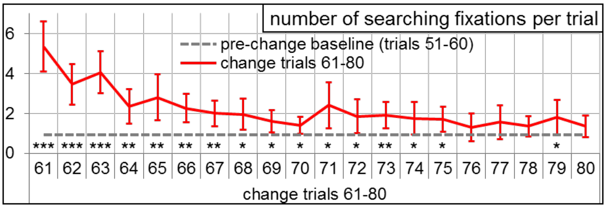

Figure 9.3: Searching fixations per change trial

Figure 9.3, which is not a part of the original published paper, shows the number of searching fixations per change trial (trials 61-80) in red solid lines along with the pre-change baseline (average of trials 51-60). The error bars represent the two-sided 95%-confidence intervals of the paired t-tests comparing the respective trial to the pre-change baseline. Asterisks indicate the two-tailored significance level (*<0.05, **<0.01, and ***<0.001).

In summary, *only the action-relevant expectation violations affected participants' manual performance and eye movements.* In this case, participants are forced to update a learned sensorimotor sequence. Thus, they *regressed from LTM-based attention and gaze control to visual search.* They maintained this search mode after having acted on the first changed number in the sequence as well as for up to 15 repetitions of the new configuration.

9.3 Analytical Reproducibility

The goal of the reproducibility experiment was to independently verify the report about performance improvements during the prechange/acquisition phase ensuring that participants adopted LTM-based attentional selection for the sensorimotor sequence. Secondly, we verified the effects of different expectation-violation manipulations on performance, eye movements and the three fixation types, allowing conclusions about the modes of attention selection, i.e. LTM versus visual search. Lastly, we verified the analysis of the repeated expectation violations updating the sensorimotor sequence based on the previously learned visuospatial task configurations which affected the mode of attentional control.

9.3.1 Research Data

The data for the entire research group are analyzed by proprietary as well as open and self-made analysis tools including SR Research's Data Viewer, SMI

BeGaze, Matlab, Python, R, SPSS, Excel, Annotation Tools [4, 15], and FuncSim [16, 17]. Experiments are programmed with SR Research's Experimental Builder, Matlab and PsychToolbox, Python and PsychoPy, E-Prime, or SMIs Experiment Suite. The data and scripts for the original work are available at `https://gitlab.ub.uni-bielefeld.de/conquaire/neurocognitive_psychology`. The folder structure is as follows:

- /loadevents
- /MatlabSkripteNFunctions
- /saveevents
- /SPSStabs
- SPSS command script
- Other project files (XLS sheets with results, etc.)

Primary Data

The data resulting from the experiment were available in a text format in the above mentioned */loadevents* folder.

Analytical Workflow

The researchers carried out their data analysis and processing in the MS Windows environment and for programming and computational analysis, they used Matlab and SPSS scripts that processed their data stored in text files. The first folder loadevents holds the data collected for each participant in six files (blinks, fixations, messages, results, saccades, and samples). Thus, the data recorded for 40 participants is held in 240 ".txt" files which became the source of input for further processing. The second folder MatlabSkripteNFunctions has 24 Matlab function scripts to perform the intermediate processing of combining and segregating data into separate event files and further input for Statistical Analysis through SPSS commands. The output of Matlab functions were stored in the third and fourth folders namely *saveevents*. The processing workflow is summarized in Figure 9.4

9.3.2 Analytical Reproducibility status

In order to reproduce the results described in the paper *'Expectation violations in sensorimotor sequences: shifting from LTM-based attentional selection to visual search'* [1], we reproduced the pipeline that was originally used to generate the results of the ANOVA and *t*-tests as described above. We could reproduce all the results as published in the original paper. For this, we acquired a 14-day

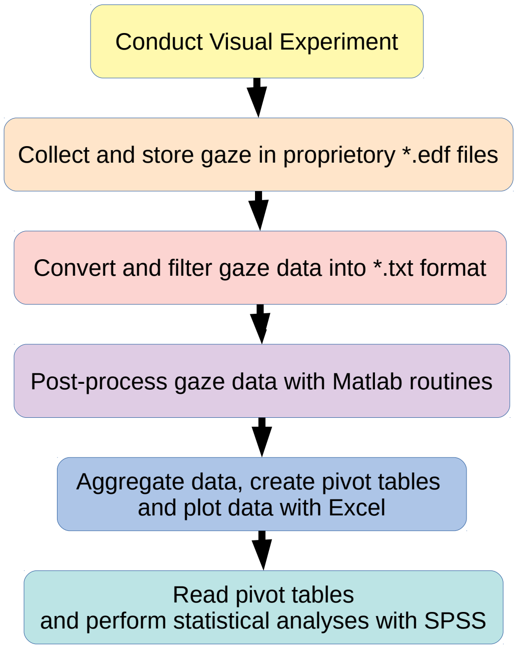

Figure 9.4: Schematic representation of the analytical workflow used in the paper by Foerster and Schneider[1]

trial version of SPSS package for the MS Windows environment from the IBM website. The SPSS script for analysis was processed to get the results published in the paper by the researchers, who confirmed that the output results were the same as the statistical results already published in the paper.

As we relied in this pipeline on proprietary and commercial software that is not freely available (Microsoft Excel, Matlab and SPSS), this is a case of limited reproducibility according to the taxonomy introduced as described in chapter 1. Thus, we also attempted to investigate whether the results could be reproduced using free and open software, in particular **Gnu-PSPP**[1] and **R**[2]. We briefly document the results of this experiment below:

SPSS vs. PSPP

PSPP was quite similar to SPSS and accepted the same code commands and in the same format as SPSS does. After making a few changes in the SPSS command file **ExpectDiscrep.sps**, the statistical tests for **NPAR TEST** and **T-TEST** ran successfully. However, other statistical functions, like **GLM TEST**, **UNIANOVA** failed as these functions are not yet implemented in PSPP. In SPSS, the GLM implements 'marginal means' but in PSPP, the **GLM** implementation is an experimental model of one-way and multiple regression linear model. So we could not reproduce the main results of the paper using PSPP.

SPSS vs. R

After investigating the use of the R-package ezANOVA, we found out that the results of the ANOVA tests could be reproduced. For the case of trend analysis, one needs to retrieve the used trends from SPSS by adding the print command 'TEST(MMATRIX)' and then insert the used contrast for the linear trend into the ezANOVA trend analysis in R. Our conclusion is thus that the results are also in principle reproducible with free and open software.

9.3.3 Discussion of reproducibility experiment

The results of the ANOVA and *t*-tests as reported in the original paper by Foerster and Schneider [1] could be independently reproduced by recreating the original analytical pipeline using the very same software stack and tools as used in the original work. This was possible because the primary data, scripts (Matlab, SPSS) and spreadsheets (Excel) were made available to the Conquaire project. Inspite of all data being available and the results being in principle reproducible, we classify this case study as an example of *limited reproducibility* as defined in the introduction to this book (see chapter 1) due to the following reasons:

[1] `https://www.gnu.org/software/pspp/`
[2] `https://www.r-project.org/`

- The analytical workflow could be reconstructed in close interaction with the authors of the original paper. The analytical workflow is not documented, so that the reproduction without guidance of the original authors is cumbersome.

- The analytical workflow relies on having installed proprietary and commercial software such as Matlab and SPSS, requiring a Windows environment for the latter. Our experiments show that substituting parts of the workflow with FOSS components, R in particular, is feasible, but this requires reprogramming the tests in R. While this is feasible, one runs the risk of creating a pipeline that is not functionally equivalent to the original one as the implementations of the tests might differ.

9.4 Conclusion

This chapter has described a case study in analytical reproducibility in the area of neuro-cognitive psychology. In particular, we have described our effort to reproduce the main results of the article by Foerster and Schneider: *'Expectation violations in sensorimotor sequences: Shifting from LTM-based attentional selection to visual search'* [1]. The main result of the article mentioned above was the finding that expectation violations in a well-learned sensorimotor sequence in humans caused a regression from LTM-based attentional selection to visual search. The authors of the original publication (also co-authors of this article) provided the Conquaire project with all primary data and all scripts and spreadsheets used to reproduce the results. While we were successful in reproducing the results, we classify this use case as one of *limited analytical reproducibility.* The reason for this is that some parts of the analytical pipeline rely on proprietary and commercial tools such as Matlab or SPSS that can not easily be replaced by open and free tools. Further, the lack of documentation of the pipeline requires interaction with the original authors to reproduce the pipeline faithfully. Both limitations could be easily overcome if further efforts are invested.

Acknowledgements

We thank Lukas Biermann and Cord Wiljes for assistance with the reproduction of the analyses.

References

[1] R.M. Foerster and W.X. Schneider. Expectation violations in sensorimotor sequences: Shifting from ltm-based attentional selection to visual search.

Annals of the New York Academy of Sciences, 1339(1):45–59, 2015.

[2] C. H. Poth, R. M. Foerster, C. Behler, U. Schwanecke, W. X. Schneider, and M. Botsch. Ultrahigh temporal resolution of visual presentation using gaming monitors and g-sync. *Behavior Research Methods*, 50(1):26–38, 2018.

[3] R. M. Foerster, C. H. Poth, C. Behler, M. Botsch, and W. X. Schneider. Using the virtual reality device oculus rift for neuropsychological assessment of visual processing capabilities. *Scientific Reports*, 6(1), 1995.

[4] R. M. Foerster, E. Carbone, H. Koesling, and W. X. Schneider. Saccadic eye movements in a high-speed bimanual stacking task: Changes of attentional control during learning and automatization. *Journal of Vision*, 11(7):1–16, 2011.

[5] M. M. Hayhoe, A. Shrivastava, R. Mruczek, and J. B. Pelz. Visual memory and motor planning in a natural task. *Journal of Vision*, 3:49–63, 2003.

[6] M. F. Land, Mennie N., and J. Rusted. The roles of vision and eye movements in the control of activities of daily living. *Perception*, 28(11):1311–1328, 1999.

[7] B. W. Land, M. F.and Tatler. Steering with the head: The visual strategy of a racing driver. *Current Biology*, 11(14):1215–1220, 2001. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11516955.

[8] M.F. Land and B.W. Tatler. *Looking and acting*. New York: Oxford University Press., 2009.

[9] J. L. Epelboim, R. M. Steinman, E. Kowler, M. Edwards, Z. Pizlo, C. J. Erkelens, and H. Collewijn. The function of visual search and memory in sequential looking tasks. *Vision Research*, 35(23-24):3401–3422, 1995. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8560808.

[10] R. M. Foerster. looking-at-nothing during sequential sensorimotor actions: Long-term memory-based eye scanning of remembered target locations. *Vision Research*, 144:29–37, 1995.

[11] R. M. Foerster and W. X. Schneider. Anticipatory eye movements in sensorimotor actions: On the role of guiding fixations during learning. *Cognitive Processing*, 16:227–231, 2015.

[12] R. M. Foerster. Task-irrelevant expectation violations in sequential manual actions: Evidence for a check-after-surprise mode of visual attention and eye-hand decoupling. *Frontiers in Psychology*, 7:1–12, 2016.

[13] G. Horstmann. Latency and duration of the action interruption in surprise. *Cognition & Emotion*, 20(2):242–273, 2006.

[14] A. Schützwohl. Surprise and schema strength. *Learning, Memory and Cognition*, 24(5):1182–1199, 1998. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9747529.

[15] R. M. Foerster, Carbone E., H. Koesling, and W. X. Schneider. Saccadic eye movements in the dark while performing an automatized sequential high-speed sensorimotor task. *Journal of Vision*, 12(2):1–15, 2012.

[16] R. M. Foerster and W. X. Schneider. Tfuncsim toolbox for matlab: Computation of eyetracking scanpath similarity, 2013.

[17] R. M. Foerster and W. X Schneider. Functionally sequenced scanpath similarity method (funcsim): Comparing and evaluating scanpath similarity based on a tasks inherent sequence of functional (action) units. *Journal of Eye Movement Research*, 6(5):1–22, 2013.