

Mouse-tracking: Detecting Types in Movement Trajectories

Dirk U. Wulff^{ab}, Jonas M. B. Haslbeck^c, Pascal J. Kieslich^d, Felix Henninger^{d,e} and Michael
Schulte-Mecklenbeck^{bf}

University of Basel, Basel, Switzerland^a

Max Planck Institute for Human Development, Berlin, Germany^b

University of Amsterdam, Amsterdam, The Netherlands^c

University of Mannheim, Mannheim, Germany^d

University of Koblenz-Landau, Landau, Germany^e

University of Bern, Bern, Switzerland^f

To appear as:

Wulff, D. U., Haslbeck, J. M. B., Kieslich, P. J., Henninger, F., Schulte-Mecklenbeck, M. (2018).
Mouse- tracking: Detecting types in movement trajectories. In M. Schulte-Mecklenbeck, A.
Kuehberger, & J. G. Johnson (Ed.), A handbook of process tracing methods (2. ed.). Psychology
Press.

Author note:

Corresponding Author: Dirk Wulff, Center for Cognitive and Decision Science, University of
Basel, Missionsstrasse 64A, CH-4055 Basel, Switzerland, dirk.wulff@unibas.ch

Keywords: Mouse-tracking, movement trajectories, cluster analysis, continuity of mind

Mouse-tracking: Detecting Types in Movement Trajectories

In recent years, mouse- and hand-tracking has become a popular method for studying the cognitive processes involved in a variety of domains, including language processing (Spivey, Grosjean, & Knoblich, 2005; Tomlinson, Bailey, & Bott, 2013), memory functions (Papesh & Goldinger, 2012; Xiao & Yamauchi, 2014), social cognition (Freeman & Ambady, 2009, 2011; Yu, Wang, Wang, & Bastin, 2012), and preferential and moral decision making (Dshemuchadse, Scherbaum, & Goschke, 2012, Koop, 2013; Koop & Johnson, 2013; Kieslich & Hilbig, 2014), to name just a few. The popularity of mouse- and hand-tracking derives from its promise to provide a window into the evolution of cognitive processes with an unrivaled temporal resolution. This window is opened by linking characteristics of the movement, such as the shape of its trajectory, to characteristics of the underlying cognitive process. In present research, this means, more often than not, inferring the response competition created by different conditions or stimuli based on the curvature of aggregate movement trajectories (see Chapters 8, 10). However, we will argue that this aggregate-level approach risks obscuring important trial-level variability in movement trajectories that can paint a different picture of the underlying cognitive process than the aggregate-level results do. Fortunately, mouse-tracking can do much more. In this chapter, we present a new approach to analyzing mouse-trajectories based on trajectory clustering that overcomes the limitations of aggregation-based analyses of movement trajectories.

Henceforth, we will focus mainly on movement trajectories recorded using the computer mouse, that is, on mouse-tracking. However the issues associated with aggregation-based analyses of movement trajectories and the usefulness of our new approach apply equally to data generated with other hand-tracking devices, such as camera-based motion-tracking (e.g., Friedman, Brown, & Finkbeiner, 2013), Wii-motes (O'Hara, Tyndall, McMorro, & Dale,

2013), touch- and track-pads (e.g., Wirth, Pfister, & Kunde, 2015), and robotic handles (e.g., Burk, Ingram, Franklin, Shadlen, & Wolpert, 2014).

Why Not Aggregate Mouse- and Hand-Trajectories?

Mouse-tracking methodologies seek to reveal the cognitive process by continuously recording the position of the computer mouse (or hand) while the subject reaches from a neutral start point to one of several spatially separated choice options. An early and very influential study that used this approach was conducted by Spivey et al. (2005). In their study, individuals were asked to use the mouse to select one of two pictures presented in the top-right and top-left corners of the screen matching a word they heard over headphones. For example, participants would see images of a ‘candle’ and a ‘candy’ and, after a short delay, be instructed to click on the image depicting the word candle (see Figure 1). To uncover how phonological competition between words is resolved in the mind, Spivey et al. (2005) implemented two conditions that varied the degree of phonological overlap between the words: In one condition, individuals chose between images depicting phonologically similar words (e.g., candle and candy; cohort condition) while in another condition they chose between images depicting phonologically dissimilar words (e.g., candle and jacket; control condition). Using this setup, Spivey et al. (2005) observed that images of phonologically similar words led to an average movement trajectory that was more drawn towards the unchosen alternative than images of phonologically dissimilar words (see Figure 1). Moreover, they found that average trajectories exhibited continuously curved trajectories. Spivey et al. interpreted these results as evidence for continuous partial co-activation of the non-chosen option that drew the movement towards the corresponding picture, expressing what Spivey later coined the *continuity of mind hypothesis* (Spivey, 2008).

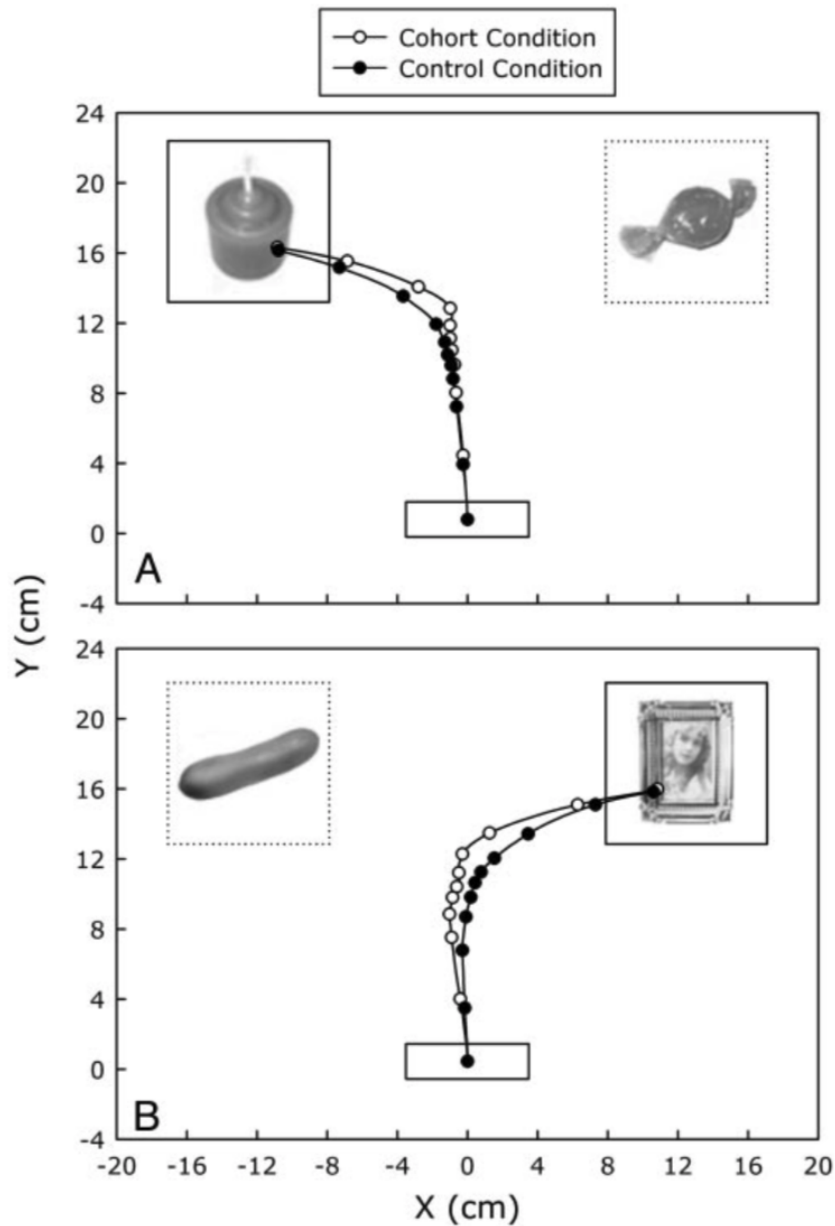


Figure 1. Design and results of Spivey et al. (2005). Lines represent the mean trajectories for the control and cohort conditions when the target was on the left (A) and right (B) side. Points are plotted every 10 normalized time slices. Reprint of Spivey et al. (2005).

Spivey et al.'s investigation built on the assumption of a close link between cognitive and motor processes on a neuronal level (Song & Nakayama, 2009). In support of this idea, Cisek and Kalaska (2005) showed that regions of the primate brain involved in planning and execution

of limb movements simultaneously encode multiple targets for selective reaching prior to choice, suggesting that “motor systems do not passively reflect the result of completed cognitive processes; rather, it is crucially linked to the dynamic decision-making process itself” (Song & Nakayama, 2009, p. 361). Movement trajectories, thus, appear to reflect the underlying cognitive process. But in what way? What exactly is the mapping between the underlying cognitive process and the movement trajectory?

Spivey et al. (2005), for instance, implicitly assumed that the underlying response competition mapped to the degree of deflection of similarly shaped curved trajectories. However, the mapping could also be of a different kind, as the authors themselves anticipated:

“In principle, it could be the case that, as with saccadic eye movements, there are some trials in which the competitor object does not attract the motor output and other trials in which it does. [...] When averaged, this hypothetical data pattern would produce mean movement trajectories that could falsely suggest simultaneous partial activation and competition [...]” (p. 10395)

In this a case, average trajectories might not necessarily represent trial-level trajectories. Instead, curved average trajectories could arise from a mixture of trajectory types, comprising both straight trajectories that immediately head to the chosen option and discrete trajectories that pass through the non-chosen option before moving to the chosen option. Average trajectories would then drastically misrepresent the trial-level trajectories in terms of both shape (no trajectory matching the continuous curve of the aggregate trajectory) and location (no trajectory spatially close to the average trajectory’s location). Moreover, differences in average trajectories could arise from small changes in the proportion of trajectory types. This means that

characteristics of aggregate trajectories could be driven by a very small number of trajectories, e.g., 5% versus 10% discrete revision trajectories.

Recent evidence has shown that the alternative mapping anticipated by Spivey et al. (2005) might represent the rule rather than the exception. Specifically, a recent reanalysis of 40 published mouse and hand-tracking studies has revealed that the majority of datasets consist of multiple types of trajectories, including the two types described in the quote above (Wulff, Haslbeck, & Schulte-Mecklenbeck, 2018). This finding has theoretical implications for what mouse-tracking reveals about the underlying process and limits the conclusions that can be drawn from smooth aggregate trajectories. It also implies that movement trajectories should be analyzed on the trial-level rather than on the aggregate-level to avoid obscuring variability in trajectory types.

Detecting Trajectory Mixtures Using Cluster Analysis

In statistics, a common approach to identifying types in a set of objects is cluster analysis (Friedman, Hastie, & Tibshirani, 2001). Cluster analysis is a class of statistical techniques that sorts objects (in our case trajectories) into a pre-defined number of clusters, choosing clusters such that the similarity of objects within each is maximized while the similarity of objects between clusters is minimized. Using cluster analysis typically requires three steps: (a) determining the distances between all pairs of objects, (b) selecting a suitable clustering algorithm, (c) choosing an appropriate number of clusters to extract. In this section, we discuss each of these steps in the context of identifying types in movement-trajectories and demonstrate the approach using the data of Spivey et al. (2005).

How to Determine Distances between Trajectories?

A fundamental challenge in using cluster analysis for mouse- and hand-trajectories is that trajectories are recorded with different numbers of points. This is because movements are typically recorded at a fixed frequency, e.g., 100Hz (100 samples per second). Trials of different duration, therefore, produce trajectories of different length, making it difficult to determine the distance between trajectories. As this problem affects many types of mouse-tracking analyses, it is common to use *time-normalized trajectories* (trajectories are interpolated such that they are represented by the same number - usually 101 - of temporally equidistant points, Spivey et al., 2005; see also Chapter 8). However, time-normalization is not the ideal choice for identifying types of trajectories: Movements are typically slow at the beginning and end of the trajectory (due to movement initiation and target approach) and fast in-between, resulting in an uneven distribution of points over the path of the trajectory. Time-normalized trajectories therefore place excessive weight on the relatively uninformative end and start phases of trajectories, where movement is slow and rarely differs in direction, and too little weight on the path in-between, where most of the movement happens and trajectories differ most from one another. A better solution is the spatial analogue of time-normalization, which represents trajectories using a fixed number of points equidistant in terms of their x- and y-positions on the screen, rather than in terms of time. Figure 2a and 2b illustrate this *spatial normalization*. The trajectory in Figure 2a represents a typical raw (or time-normalized) trajectory that has most points in the beginning and the end of the trajectory. By contrast, the points of the spatially normalized trajectory in Figure 2b are equally distributed in the 2D plane. As a result, the spatially normalized trajectories are represented with uniform precision, giving every part of the trajectories' shapes equal weight.

A key parameter of space-normalization is the number of points used to represent the trajectory. Traditional analyses normalize trajectories to 100 or 101 points. For identifying types, however, it is practical to space-normalize trajectories to a substantially smaller number of points, e.g., 20. The first reason for this is that the clustering procedure calculates the distance for every pair of trajectories, which can become computationally expensive if both the number of trajectories and the number of points per trajectory is large. Second, choosing a relatively small number of points to represent trajectories places the focus on the overall shape of the trajectory rather than its details, which aids the detection of commonalities between trajectories and, thus, the identification of basic types.

Once trajectories are represented by equally many points, calculating distances between trajectories is straightforward using, for instance, the Euclidean distance (see Figure 2c). The Euclidean distance d_{ij} between two trajectories i, j is given by

$$d_{ij}(x_i, y_i, x_j, y_j) = \sqrt{\sum_k (x_{ik} - x_{jk})^2 + (y_{ik} - y_{jk})^2} \quad (1)$$

where x and y are the x - and y -coordinates of trajectories i and j . However, other distance metrics are possible (e.g., city block) and justifiable. Having decided on a distance metric, the next step is to determine the full distance matrix by computing the distance between every pair of trajectories in the data set (Figure 2d). This step can be very computationally expensive, as the needed amount of computation grows quadratically with the number of trajectories. This problem can be alleviated by using the high-performance tools implemented in the mousetrap R package (see Chapter 8).

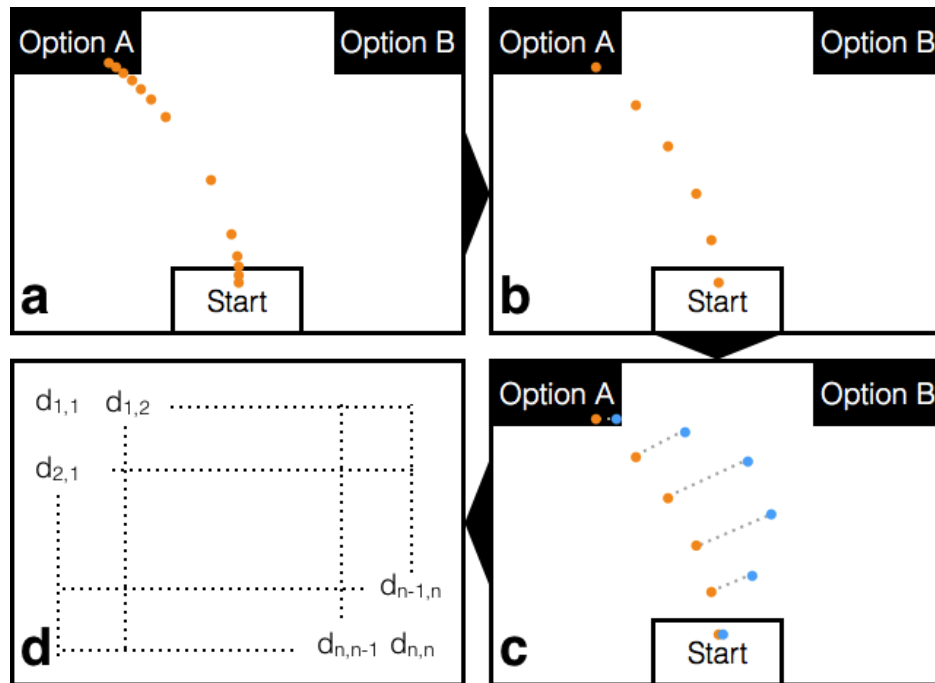


Figure 2. Computation of pairwise trajectory distances. Panel (a) illustrates the distribution of trajectory points in an exemplary raw (or time-normalized) trajectory. Panel (b) illustrates how the spatial normalization approach distributes trajectory points equally across the 2D-plane to give relatively more weight to the trajectories' body. Panel (c) illustrates the difference between two exemplary spatially normalized trajectories that give rise to one element in the distance matrix illustrated in panel (d).

How to Cluster Trajectories?

Having obtained the distance matrix, the trajectory data is ready for use for many off-the-shelf cluster analysis algorithms, which commonly take a distance matrix as input. Two of the most popular clustering algorithms are hierarchical clustering and k-means clustering (Friedman, Hastie, & Tibshirani, 2001): Hierarchical clustering builds a hierarchy of clusters by either

joining objects from the bottom up (agglomerative clustering) or splitting clusters from the top down (divisive clustering) using one of many available linkage criteria¹. Single-linkage, agglomerative clustering criterion, for instance, will begin with each object occupying its own cluster and then proceed by combining at each step two clusters that contain the closest pair of elements. *K*-means clustering, on the other hand, partitions the objects into *k* clusters by iteratively assigning an object to the cluster with the nearest cluster centroid, as measured by their Euclidian distance. Distinguishing the two algorithms is mainly the criterion used to form clusters. Whereas *k*-means is based on Euclidian distance and minimizing within-cluster variance, hierarchical clustering offers more flexibility. However, we have found that, in the case of movement trajectories, hierarchical clustering also performs best using criteria that minimize within cluster variance (e.g., Ward's *d*).

How Many Clusters to Extract?

A final consideration in extracting clusters is to choose the number of clusters *k*. Practically all cluster analysis techniques require *k* to be specified by the user. Unfortunately, there exists a fundamental problem that in one way or another pertains to all unsupervised learning algorithms: there is no principled way to evaluate the quality of a cluster solution given a certain *k*. This would only be possible if we knew the data generating process, which would render clustering unnecessary. There are, nonetheless, several methods to estimate the number of clusters each based on a different criterion of a good clustering (see Henning, 2015, for a discussion). One such method is *k*-selection based on cluster stability (see e.g., Ben-Hur, Elisseeff, & Guyon,

¹ Agglomerative clustering is usually preferred as it shows much lower computational complexity than divisive clustering.

2001; Haslbeck & Wulff, 2016; Tibshirani & Walther, 2005). This method seeks to identify the k that leads to most stable clustering in terms of its robustness against perturbations of the data. Other popular k -selection methods are the Gap statistic (Tibshirani, Walther, & Hastie, 2001), the Jump statistic (Sugar & James, 2011), and the Slope statistic (Fujita, Takahashi, & Patriota, 2014). As these methods rest on different definitions of a good clustering they often lead to different k estimates. In fact, when implementing the four methods mentioned above using the *cstab* package² (Haslbeck & Wulff, 2018) for the data sets contained in our meta-analysis the k -estimates varied widely (between 2 and 20 clusters, with 20 being the maximum number of clusters considered) and the correlations between k -estimates based on the four methods were negative across data sets. This low level of consistency between the methods illustrates the fundamental problem associated with identifying the true number of clusters.

However, the situation is actually not that dire. If the data is composed of clearly distinguishable types, then 2D-plots of the extracted clusters should show qualitatively distinct trajectory shapes. Moreover, it is in fact not completely necessary to have identified the true set of types underlying a data set or know their exact number. Indeed, we argue that in the presence of some clearly identifiable types it always will be better to analyze the data using a type-based approach than to resort to averaging trajectories.

² The *cstab* package can also be accessed via the mousetrap wrapper function *mt_cluster_k*.

Clustering the Data of Spivey et al. (2005)

To demonstrate our approach, we reanalyzed the data of the experiment of Spivey et al. (2005)³ presented earlier. To do this, we first aligned the start points of all trajectories and flipped all trajectories that ended on the right option to the left side. Next, we space-normalized trajectories to 20 points each. Finally, we extracted four clusters⁴ using the *k*-means algorithm separately for the cohort and control condition. These and all other analyses in this chapter were run using the mousetrap R package (see Chapter 8; Kieslich et al., 2018).

This analysis revealed four important insights. First, the four clusters extracted for the cohort condition (Figure 3, top row) were remarkably similar to the four clusters extracted for the control condition (Figure 3, bottom row), speaking to the reliability of behavior and the analysis approach. Second, the analysis produced clusters of high homogeneity for both conditions, suggesting the presence of distinct types: A first cluster of trajectories that go straight up to the level of the two options and then directly to the chosen option, a second cluster of trajectories that head straight to the chosen option, a third cluster of trajectories that go from the start to the chosen option in a mild curve, and the fourth and final cluster whose trajectories visit the non-chosen option before going to the chosen option. Third, the clustered types of trajectories occurred in different proportions between conditions. In particular, the conditions differed mainly in terms of trajectories that go straight to the chosen option (23% versus 33%) and trajectories that visit the non-chosen option first (20% versus 8%). Fourth, the majority of

³ To whom we are grateful for providing us with the raw data.

⁴ We arbitrarily chose to extract four clusters, as this number appeared large enough to tease apart substantive clusters and small enough to not fraction. Clusters were extracted using the *mt_cluster* function of the mousetrap R package.

trajectories (all but the third cluster) showed a shape that does not match average the trajectories presented in Figure 1.

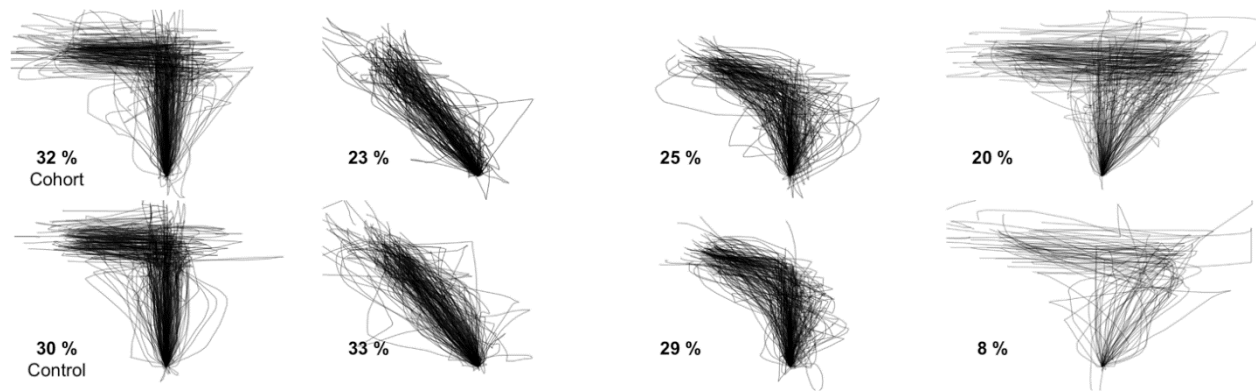


Figure 3. Trajectory clusters extracted from the data of Spivey et al. (2005). Figure shows in descending order of size (determined based on both conditions) the trajectory clusters extracted for the cohort (top-row) and control (bottom-row) condition. Percentages indicate the proportion of trajectories within a condition assigned to the respective cluster.

The above results strongly suggest the presence of different trajectory types. These types are, in three out of four cases, not well represented by the aggregate trajectory, and their relative proportions appear to produce the effect visible in Figure 1. This pattern of results has important consequences for how movement trajectories should be analyzed and interpreted. However, before we turn to these issues, let us address two important caveats.

First, to reiterate a point made earlier, clusters need not indicate types. That is, cluster analysis algorithms always produce as many clusters as they are instructed to produce. Furthermore, the extracted clusters will always contain non-overlapping sets of objects. It is, thus, possible that the clusters originate from a single cluster of trajectories that was arbitrarily divided. In fact, as mentioned above, there is no way of knowing for sure whether the data is truthfully composed of different clusters or not. Still, we can make inferences about the presence or

absence and the number of clusters. Consider a situation in which data originally came from a single cluster. In that case, trajectories should come in highly similar shapes. For the present analysis, this is however clearly not the case (Figure 3): For instance, the trajectories of two clusters show a clear kink, either at the point between the options or at the location of the non-chosen option, whereas the trajectories of the other two clusters do not. Thus, in lieu of a principled statistical technique to decide on the number of clusters, the visual inspection of clusters reveals clear qualitative differences between movement trajectories speaking in favor of trajectory types.

The second caveat is that, despite the high similarity across conditions, the clusters of the conditions are not identical. It is, therefore, difficult to compare cluster sizes across conditions and to characterize the effect of, e.g., a manipulation like the one used by Spivey et al. (2005; Figure 1) in terms of trajectory types. To do this, it would be necessary to keep clusters constant across conditions. The next section describes such an approach based on prototype matching.

Detecting Trajectory Mixtures Using Prototype Matching

Cluster algorithms practically never produce the exact same partitioning of the object space into clusters for different datasets. This makes it difficult to compare cluster solutions across different datasets or different experimental conditions. One way to address this issue is to define prototypical trajectories, and then to assign each trajectory to the closest prototypes based on some distance metric.⁵ This is equivalent to creating a fixed partitioning of the object space.

⁵ In the case of a Euclidian distance (defined in (1)), this essentially means to define in advance the cluster centroids that k -means would otherwise have to learn from the data.

As a result, clusters belonging to a given prototype will always contain the same kinds of trajectories, rendering the clustering of the data sets immediately comparable.

Another benefit of this approach is that one can tease apart qualitatively distinct trajectories that statistical algorithms may have difficulty separating. For instance, one might argue that the fourth cluster in Figure 3 is actually composed of two distinct trajectories: those that go straight up to the level of the options and then visit the non-chosen before going to the chosen option and those that go straight to the non-chosen option before going to the chosen option. Using a prototype-matching approach, this can be easily dealt with by defining separate prototypes for these sets of trajectories. In the case of Spivey et al., it might, thus, be reasonable to map trajectories onto five prototypes, three matching the first three clusters types in Figure 3 and two for the final cluster. For other data sets obtained using different experimental designs, however, different prototypes can be required.

Figure 4 shows these prototypes (Figure 4a), as well as the trajectories of the cohort (Figure 4b) and control (Figure 4c) that have been mapped to the closest prototype using Euclidian distance as defined in (1).⁶ It can be seen that the extracted trajectory clusters represent much cleaner types than the clusters extracted using k-means. This is particularly the case for the fourth cluster of the k-means solution (Figure 3), which is now separated into two clusters that clearly represent the corresponding prototype. However, an improved differentiation of types can also be observed for the other three clusters.

⁶ Prototype-mapping was carried out using the *mt_map* function of the mousetrap R package.

Based on the prototype mapping, we can now characterize the effect of the manipulation of Spivey et al. (2005) by comparing the proportions of trajectories mapped to each prototype. This reveals a similar pattern of differences to that obtained using the k-means clustering: The cohort condition included more trajectories of types four and five, and thus of trajectories that visit the non-chosen option, whereas the control condition contained more straight trajectories of type two. Thus, the difference between the conditions in aggregate curvature arises mainly from trajectories that are extremely deflected and trajectories that are not deflected at all. Next, we will discuss what conclusions to draw from this and what this means for the interpretation of mouse-tracking to study cognition.

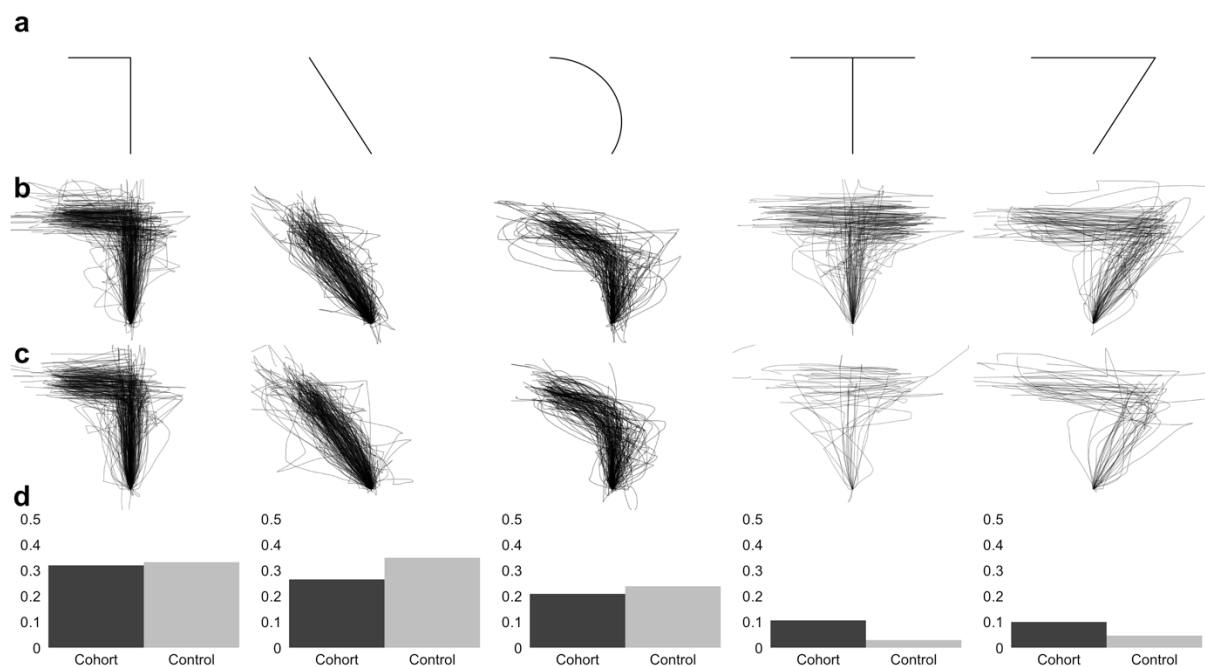


Figure 4. Prototype-matching of the data of Spivey et al. (2005). Panel (a) shows the five prototype trajectories. Panel (b) and (c) show the trajectories of the cohort and control conditions, respectively, separately for the prototypes they have been matched to. Panel (d) shows the proportions of trajectories in the cohort (dark grey) and control (light grey) conditions matched to each of the prototypes.

Implications for Mouse- and Hand-Tracking as a Window to Cognition

Our reanalysis of the data of Spivey et al. (2005) has produced a set of at least five qualitatively distinct types of trajectories. This finding raises several questions regarding the analysis of movement trajectories and what can be inferred about the underlying cognitive process from trajectory shapes.

How Prevalent are Trajectory Types?

One fundamental question is whether the types obtained for the data of Spivey et al. (2005) generalize to other data sets of mouse- and hand-tracking. Based on our meta-analysis of 40 data sets the answer is: it depends (Wulff, Haslbeck, & Schulte-Mecklenbeck, 2018). The majority of data sets, in particular those collected using the computer mouse as input device exhibited trajectories similar to those found in the data of Spivey et al. (2005). However, several data sets, particularly those collected with cameras (e.g., Friedman, Brown, & Finkenbeiner, 2013) or a robotic handle (e.g., Burk et al., 2014) showed relatively homogenous sets of smoothly curved trajectories that might be best described by a single cluster. Other important moderators are characteristics of the study design (for discussions, see Kieslich & Henninger, 2017; Kieslich, Schoemann, Grage, Hepp, & Scherbaum, 2018; Scherbaum & Kieslich, in press; Wulff, Haslbeck, & Schulte-Mecklenbeck, 2018). Figure 5 gives an illustration of the trajectories produced by four hand-picked studies, including the study by Spivey et al. (2005; Figure 5a), a mouse-tracking study using a similar methodological setup as Spivey et al. (Koop & Johnson, 2013; Figure 5b), a mouse-tracking study using a different methodological setup than Spivey et al. (Scherbaum et al., 2010; Figure 5c), and a hand-tracking study using a robotic handle (Burk et al., 2014; Figure 5d). As can be seen, not all studies tended to produce the types found in the data of Spivey et al. (2005) and some may not contain qualitatively different types at all. This implies

that mouse-tracking researchers need a better understanding about how the methodological setup influences trajectory shapes and that they should investigate the nature of trajectories in each individual study before choosing a particular path of analysis.

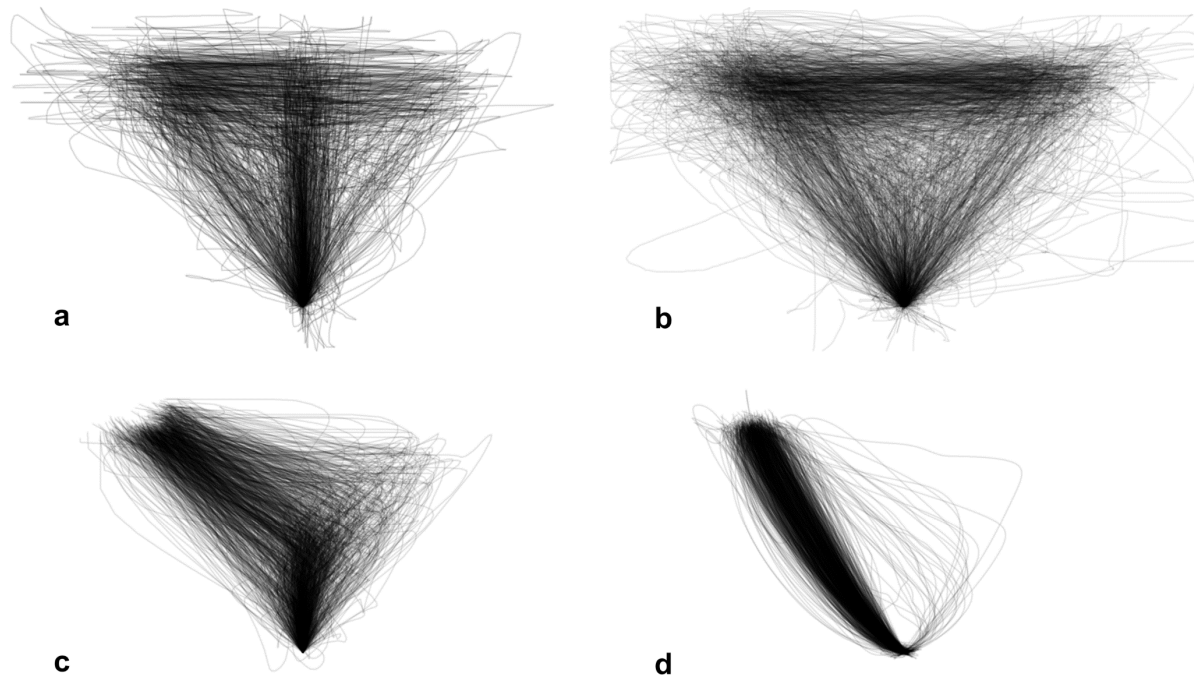


Figure 5. Overview of trajectories produced by different studies. Panel a shows the trajectories of the cohort condition of Spivey et al. (2005). Panel b shows the trajectories of choice of the risky option in experiment 2 of Koop and Johnson (2013), who used a similar design as Spivey et al. (2005). Panel c shows the trajectories of incongruent trials in Scherbaum et al. (2010), who used a different design, in which participants had initiate their movement first in order to trigger stimulus presentation. Panel d shows the trajectories of trials with narrow option separation in Burk et al. (2014), who tracked the movement of the hand using a robotic handle.

One way the literature has previously attempted to test for the presence of trajectory types is via bimodality tests. Bimodality tests rest on the idea that trajectories come in two types, deflected and not-deflected, which is outlined in the quote of Spivey et al. (2005) above.

Presence of these two types should result in a bimodal distribution of curvature measures. In that case, so the logic, measures and tests of bimodality, such as the bimodality coefficient (SAS Institute, 1989) and Hartigan's dip-statistic (Hartigan & Hartigan, 1985), would be able to identify the presence of (two) trajectory types. There are, however, two important problems with this approach. First, bimodality tests require that each trajectory is condensed to a single curvature value, rendering indistinguishable types with identical curvature values (e.g., in terms of MAD; see Chapter 8), such as the fourth and fifth prototype in Figure 4. Second, bimodality tests are specialized for detecting the presence of two modes. As soon as more trajectories are involved, bimodality tests cannot be expected to reliably identify the presence of different types. As the number of types is usually not known upfront, bimodality tests do not provide a reliable answer regarding the presence or absence of different trajectory types. A generally more suited approach to evaluate for the presence of types is to employ a bottom-up clustering approach and to visually inspect plots of raw trajectories by cluster.

How to Conduct Statistical Analyses using Trajectory Types?

Another question is how one is to conduct group comparisons or statistical analyses, in general, when the trajectory data comes in types. Typically, analyses of mouse- and hand-trajectories are based on continuous trial-level statistics. That is, trajectories are condensed into single values representing, for instance, the trajectories' curvature, their complexity, or their velocity (see Chapter 8). Such analyses are sensible when the data stems from a homogenous distribution of trajectories that vary with respect to the statistic of choice. However, when trajectories occur in different types, several problems may ensue. For instance, the presence of types may imply a multi-modal distribution of, e.g., curvature, which cannot reliably be summarized by means of a central tendency, and which violates normality assumptions made by many statistical tests. More

importantly, condensing trajectories to a single value may obscure important variability in trajectories as exemplified by cluster four and five in Figure 4 that may offer valuable insight into the underlying cognitive process.

A more appropriate statistical approach is to analyze trajectory types as a categorical variable. For instance, researchers can test the equality of trajectory distributions across different conditions using chi-square tests of independence. For the data of Spivey et al. (2005), this approach revealed a systematic difference between the trajectories produced by the two conditions. Analyzing the residuals furthermore revealed that this difference arises mainly from the types two, five, and, in particular, four (see Table 1). Analyzing the data on the nominal delivered insights not readily available otherwise, such as which trajectory types were most affected by experimental manipulations, potentially opening a new window into the underlying cognitive process.

Table 1

Type Frequencies, Proportions, and Standardized Residuals for the Data of Spivey et al. (2005)

Condition	Types				
	1	2	3	4	5
<i>Frequencies</i>					
Cohort	215 (32%)	178 (26%)	141 (21%)	71 (11%)	67 (10%)
Control	207 (33%)	218 (35%)	149 (24%)	19 (3%)	29 (5%)
<i>Residuals</i>					
Cohort	-0.28	-1.93	-0.78	3.55	2.43
Control	0.29	2.00	0.81	-3.69	-2.52

Note. Standardized residuals are based on a test of stochastic independence between the type-frequency distributions of the cohort and control conditions in Spivey et al. (2005).

Another possibility for analyzing trajectory types is to order types with regard to a dimension of interest and then to run analyses using the resulting ordinal variable. Mouse-tracking studies are, for instance, often interested in evaluating differences in curvature as a signal of the degree competition between the response option. Such an analysis can also be carried out using trajectory types, by ordering the clusters (based on, e.g., *k*-means or prototype mapping) with regard to the degree of response competition that they embody and using statistical tools suited for ordinal data, such as ordinal logistic regression. In the case of the types extracted for the data collected by Spivey et al. (2005; Figure 4), one may argue that an appropriate ordering of the types from low to high competition is cluster 2 < cluster 3 < cluster 1 < cluster 4 < cluster 5. Using this ordering, we again found that the manipulation of phonological similarity had an effect on the distribution of trajectories, this time confirming directly that the cohort condition produced trajectories consistent with higher degrees of competition.

How to Interpret Trajectory Types?

Trajectory types offer the researcher interesting new possibilities for the analysis of movement trajectories. What we have not addressed sufficiently, so far, is how to interpret results of type-based analyses on the level of cognitive processes. That is, which conclusions regarding the underlying process can be drawn from, for instance, the finding that Spivey et al.'s (2005) manipulation affected only specific kinds of trajectories, and how can this be reconciled with the typical interpretation of movement trajectories. For Spivey et al. (2005) and large parts of the mouse-tracking literature, (curved) trajectories are thought to arise from a continuous, partial co-

activation of the available options, expressed continuously in the movement. However, this is not easy to reconcile with that fact that the trajectories impacted by Spivey et al.'s (2005) manipulation, i.e., those assigned to prototypes two, four, and five, either are perfectly straight or exhibit characteristic kinks in the trajectory shape. Such trajectories suggest either (a) that the cognitive process is inherently discrete or (b) that the underlying process expresses itself on a much coarser time-scale than typically assumed.

Currently, too little is known about the link between the cognitive process, the movement trajectory, and the design that elicits the movements to discern whether (a), (b), or potentially both are the case. It may seem plausible that the kinks in the trajectories result from a non-continuous mapping of the underlying process to the movement, considering that cognitive processes reside in the near-continuous communication of neurons in the brain (Cisek & Kalaska, 2005; Song & Nakayama, 2009). However, decision scientists have also long been discussing lexicographic decision processes that consider pieces of information in sequence and thus could account for preference shifts as embodied by trajectory types four and five (e.g., Payne, Bettman, & Johnson, 1993). For the usefulness of mouse-tracking as a window to the underlying cognitive process the latter scenario would be more desirable. That is, if kinks in trajectories arise because of a low-resolution mapping of the movement to the underlying cognitive process, then movements provide only a very narrow window into the development of the cognitive processes. However, if those trajectories arise, because of inherently discrete cognitive processes, then movements may reveal not only the degree of competition between choice options, but also the discrete cognitive steps that let up to the decision.

Future research needs to study the link between movement trajectories and cognitive processes and build theories to describe the cognitive process and its link to movement

trajectories. Promising efforts in this direction have been made based on sequential-sampling models (e.g., Resulaj, Kiani, Wolpert, & Shadlen, 2009; Wong & Haith, 2002), accumulation models (e.g., Frisch, Dshemuchadse, Görner, Goschke, & Scherbaum, 2015), and decision landscapes (e.g., Zgonnikov, Aleni, Piiroinen, O'Hora, & di Bernardo, 2017). However, the empirical results on the validity of these models are still outstanding, including, in particular, the question whether these models can account for the trajectory types presented here and the existence of trajectory types, in general.

Conclusion

Mouse-tracking is a promising new tool in the study of cognitive processes. This tool is not and, as we have argued, should not be limited to the aggregate-level analyses of movement trajectories still commonly employed in the literature. Instead researchers should consider analyzing trajectories on the trial-level using, for instance, the clustering approaches presented here and made available through the mousetrap R package. Frequently this will reveal that movement trajectories occur in distinct trajectory types, challenging the common conception of movement trajectories as a direct window to continuous co-activation of choices options. Further research is needed to better understand the link between the cognitive processes and movement trajectories, and how it is best utilized to reveal the underlying cognitive processes. To that end, the analysis of trajectory types presents an exciting new avenue offering rich data ready to be explored.

Recommended reading list

- Spivey, Grosjean, & Knoblich (2005): Early and highly influential study demonstrating a link between response competition and trajectory curvature.
- Haslbeck, & Wulff (2016): Formal study discussing different methods to estimate the number of clusters in empirical data sets.

References

- Ben-Hur, A., Elisseeff, A., & Guyon, I. (2001). A stability based method for discovering structure in clustered data. *Pacific Symposium on Biocomputing*, 7, 6–17.
- Burk, D., Ingram, J. N., Franklin, D. W., Shadlen, M. N., & Wolpert, D. M. (2014). Motor effort alters changes of mind in sensorimotor decision making. *PLoS ONE*, 9(3), 92681.
- Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron*, 45(5), 801–814.
- Dshemuchadse, M., Scherbaum, S., & Goschke, T. (2013). How decisions emerge: Action dynamics in intertemporal decision making. *Journal of Experimental Psychology: General*, 142(1), 93–100.
- Freeman, J. B., & Ambady, N. (2009). Motions of the hand expose the partial and parallel activation of stereotypes. *Psychological Science*, 20(10), 1183–1188.
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, 118(2), 247–279.
- Friedman, J., Brown, S., & Finkbeiner, M. (2013). Linking cognitive and reaching trajectories via intermittent movement control. *Journal of Mathematical Psychology*, 57(3–4), 140–151.
- Friedman, J. H., Hastie, T., & Tibshirani, R. (2001). *Springer series in statistics: Vol. 1. The elements of statistical learning*. New York, NY: Springer.

- Frisch, S., Dshemuchadse, M., Görner, M., Goschke, T., & Scherbaum, S. (2015). Unraveling the sub-processes of selective attention: Insights from dynamic modeling and continuous behavior. *Cognitive Processing*, 16(4), 377–388.
- Fujita, A., Takahashi, D. Y., & Patriota, A. G. (2014). A non-parametric method to estimate the number of clusters. *Computational Statistics & Data Analysis*, 73, 27–39.
- Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*, 13(1), 70–84.
- Haslbeck, J., & Wulff, D. U. (2016). Estimating the number of clusters via normalized cluster Instability. *arXiv:1608.07494*.
- Kieslich, P. J., & Hilbig, B. E. (2014). Cognitive conflict in social dilemmas: An analysis of response dynamics. *Judgment and Decision Making*, 9(6), 510–522.
- Kieslich, P. J., & Henninger, F. (2017). Mousetrap: An integrated, open-source mouse-tracking package. *Behavior Research Methods*, 49(5), 1652–1667.
- Kieslich, P. J., Schoemann, M., Grage, T., Hepp, J., & Scherbaum, S. (2018). *Design factors in mouse-tracking: What makes a difference?* Manuscript in preparation.
- Kieslich, P. J., Wulff, D. U., Henninger, F., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2018). *Mouse- and hand-tracking as a window to cognition: A tutorial on implementation, analysis, and visualization*. Manuscript in preparation.
- Koop, G. J., & Johnson, J. G. (2013). The response dynamics of preferential choice. *Cognitive Psychology*, 67(4), 151–185.
- Koop, G. J. (2013). An assessment of the temporal dynamics of moral decisions. *Judgment and Decision Making*, 8(5), 527–539.

- O'Hara, D. P., Tyndall, I. T., McMorrow, M., & Dale, R. A. (2013). Using action dynamics to assess competing stimulus control during stimulus equivalence testing. *Learning & Behavior*, 41(3), 256–270.
- Papesh, M. H., & Goldinger, S. D. (2012). Memory in motion: Movement dynamics reveal memory strength. *Psychonomic Bulletin & Review*, 19(5), 906–913.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge, UK: Cambridge University Press.
- Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N., (2009) Changes of mind in decision-making. *Nature*, 461(7261), 263–266.
- SAS Institute Inc. (1989). *User's guide*. Cary, NC: Statistical Analysis System Institute.
- Scherbaum, S., & Kieslich, P. J. (in press). Stuck at the starting line: How the starting procedure influences mouse-tracking data. *Behavior Research Methods*, 1–14.
- Scherbaum, S., Dshemuchadse, M., Fischer, R., & Goschke, T. (2010). How decisions evolve: The temporal dynamics of action selection. *Cognition*, 115(3), 407–416.
- Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, 13(8), 360–366.
- Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences of the United States of America*, 102(29), 10393–10398.
- Spivey, M. (2008). *The continuity of mind*. Oxford, UK: Oxford University Press.
- Sugar, C. A., & James, G. M. (2003). Finding the number of clusters in a dataset: An information-theoretic approach. *Journal of the American Statistical Association*, 98(463), 750–763.

- Tibshirani, R., & Walther, G. (2005). Cluster validation by prediction strength. *Journal of Computational and Graphical Statistics*, 14(3), 511–528.
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411–423.
- Tomlinson Jr, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, 69(1), 18–35.
- Wirth, R., Pfister, R., & Kunde, W. (2016). Asymmetric transfer effects between cognitive and affective task disturbances. *Cognition and Emotion*, 30(3), 399–416.
- Wong, A. L., Haith, A. M., & Krakauer, J. W. (2015). Motor planning. *The Neuroscientist*, 21(4), 385–398.
- Wulff, D. U., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2018). *Measuring the (dis-)continuous mind: What movement trajectories reveal about cognition*. Manuscript in preparation.
- Xiao, K., & Yamauchi, T. (2014). Semantic priming revealed by mouse movement trajectories. *Consciousness and Cognition*, 27, 42–52.
- Yu, Z., Wang, F., Wang, D., & Bastin, M. (2012). Beyond reaction times: Incorporating mouse-tracking measures into the implicit association test to examine its underlying process. *Social Cognition*, 30(3), 289–306.
- Zgonnikov, A., Aleni, A., Piironen, P. T., O'Hora, D., & di Bernardo, M. (2017). Decision landscapes: Visualizing mouse-tracking data. *Royal Society Open Science*, 4(11), 170482.