Discover the Invisible Through Tool-Supported Scientific Observation

A Best Practice Guide to Video-Supported Behavior Observation

Pascal Mangold

Observation appears to be a simple skill. It is assumed to be something everyone does every day since early childhood, and thus, it seems to be an easy and well-trained skill. However, observation of behavior with a scientific outcome is far from easy. It requires a well-thought-out method based on scientific knowledge and hypothesis. It further requires appropriate software tools to create reliable data and new findings with significant validity in a reasonable time. The significant difference between everyday observation and scientific observation and the enormous chances specific software tools can create in this field will be discussed in this document.

Keywords: scientific observation, video technology, video rating software, Mangold INTERACT

1 Motivation

This paper presents the key aspects of successful scientific observational studies. It shows the added value that can only be achieved with the support of a specialized observational software tool and clarifies the necessary framework conditions for the most effective and efficient observation.

2 Everyday Observation and Scientific Observation

Observing is a word in common parlance. It is used for sentences like: "I have observed that...", "Keep this under observation", "Our observations are based on...". Such phrases imply that all participants have the same understanding of what observation is. However, this is often not the case, as it can easily be verified in everyday life when expectations and results do not match. Therefore, it seems necessary to first define the term "observation" as such. An apt definition is:

Definition 1: Observation is the "purposeful, attentive perception of objects, phenomena or processes, if necessary, using technical aids" (FH Münster 2017).

A distinction must be made between everyday observation and scientific observation:

For humans, everyday observation essentially takes place at any time. It is arbitrary, driven by our expectations, and directed towards our needs. Nevertheless, it can also be completely random. It can be essential to life, for example, when carrying out activities that are necessary daily. It can even be existential, for example, when operating dangerous machines or crossing a busy road. In such cases, everyday observations and possible conclusions must be made very quickly. Therefore, intuition and subjective assessment are used as decision criteria to achieve a high processing speed. Since everyday observations are under certain circumstances driven by the compulsion to make a quick decision, we call them result- driven[1].

The scientific observation, on the other hand, is clearly different. As a data collection method in scientific studies, observation is necessary because it supplements other methods with information that cannot be collected, for example, by measurements or surveys for various reasons (cf. McLeod 2015; Faßnacht 2007; Greve & Wentura 1997). In contrast to everyday observation, it is not arbitrary but more or less strongly structured, depending on the object of observation. However, it cannot be completely structured because it would have to be possible to strictly define all aspects of the observation. If this were the case, these aspects could be represented in technical structures (software/hardware) so that the whole process would no longer be an observation but a pure measurement. The possibility to proceed less stringently is a tremendous advantage of observation itself, but it also holds considerable potential for errors in the collected data[2]. To ensure a structured approach, observation needs to be operationalized. For this purpose, the aspects and contents to be observed must be clearly described in advance (more on this later). The scientific observation is thus no longer intuitive and subjective, like the everyday observation, but as objective as possible[3]. It is value-free, purely descriptive, and abstract. It is fundamentally open-ended compared to everyday observation since it aims to generate data as precisely as possible. This definition allows a clear separation of data collection and the subsequent data processing, which gives these data a meaning. This is particularly helpful since the data do not incriminate subliminal information with them but can be processed in a purely abstract way[4].

For the success of this data collection method, some essential aspects must be considered. (FH Münster 2017) lists the following elements as main quality criteria:

[1]Not always, because letting one's thoughts wander with simultaneous observation of the passing clouds on a beautiful vacation day can also be defined as everyday observation. It is, like the chance observation, not result driven.
[2] Also, the waste of time is a big problem with less structured observations. This time is usually missing elsewhere in a study, which can ultimately affect the quantity and/or quality of the results.
[3] A complete objectification would be equivalent to a measurement that differs from the observation. A high degree of objectivation is of course desirable to have as little error and room for interpretation in the results as possible.
[4] Of course, this enables an easily generated misleading correlation between the collected data. But in any case, it is helpful not to limit the possible result space by all kinds of restrictions from the beginning, but to eliminate nonsensical artifacts from the result space during the data evaluation. Something new can rather emerge if it is not limited in advance by the space of the supposedly known.

- Objectivity (as independence from researcher and result)
- Reliability (of, e.g., the data collection method) and
- Validity (as proof that what you wanted to capture has been captured).

These aspects are undoubtedly decisive for the result, whereby the following points are essential on the way to finding the result:

- The content-related aspects of the observation (what is the reason for the observation).
- The observation procedure (what and how to observe).
- The methods of gaining knowledge (how to transform the data collected by observation into meaningful results).

Since scientific observation is the subject of this document, and everyday observation was only presented as a conceptual delimitation, the term “observation” continues to refer to the scientific form.

3 Motivation and contents of an observation

If one asks why anything should be observed at all, it quickly becomes clear that observation is fundamental to all scientific activities and all scientific disciplines (Norris 1984). Observation helps to understand any processes whereby a process can be defined as follows:

Definition 2: “A process is a flow of interrelated activities that lead to a known or unknown result”. If this definition is divided into its components, the following results:

“Activity”: An activity is something clearly describable, such as “Person A gives person B the object X”.

“Relation”: The relation indicates how and which activities are related to each other. In the example above, person B can “take” or “reject” the object. However, B can only do this if the object was offered before. Therefore, the activities “take” and “reject” are dependent on the predecessor.

A common form of conducting observational studies is the more or less structured recording of such activities: For example, ad-hoc, by written live protocols, in the form of live filled out checklists or by post-hoc recording, using structured questionnaires or free memory protocols.

The definition of such activities and the expectation of a connection already allow to work out results. These can be *known* or *expected results* (e.g., “offer” is followed by “take” or “reject”) but also *unexpected results* like “object falls”, or “object bends and cannot be handed over"[5].

The possible result space in which the activities are related is two-dimensional, as shown in Figure 1. One could criticize that also several activities can occur simultaneously, which in turn are related to any number of simultaneous activities, making the space multi-dimensional (e.g., "verbal offering and simultaneous pointing to the object" is related to "the object is taken with simultaneous thanks"). This view of the two dimensions is simplified but correct since the definition of actions is merely a matter of definition, depending on the observation's content. Thus, the combination of several behaviors can again be seen as one (complex) action. This action can be understood as a vector of different objects or phenomena, e.g., Action "A" = (verbal offer + pointing gesture), "B" = (verbal thanks + physical taking). Seen in this way, A and B are in an identical relationship, like the detailed form shown above with four actions (offer, show, take, thank). A problem with observational studies arises when the "flow" of activities is recorded inaccurately or not at all. It is precisely this captured flow in the form of time information that carries essential information. This information forms an additional, content-wise dimension in the space of the possible results. The decisive factor here is that time is recorded as accurately as possible (depending on the observation subject, accurate[6] to milliseconds). As shown in Figure 2, the correlations' space to be investigated, and consequently, the number of expected results is now much higher. Especially if not only pure sequences of activities are to be investigated, but also their change over time (for example, in therapies, during feedback conversations, or in case of changes due to learning behavior), the temporal flow is one, if not the crucial criterion[7] for interesting results.

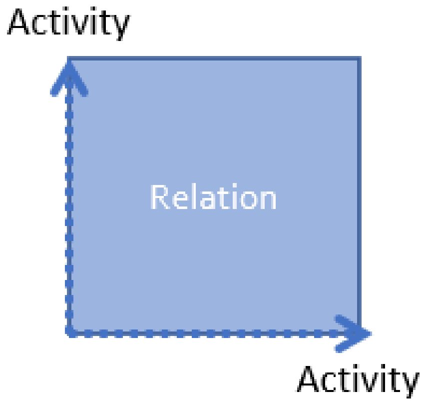

Figure 1

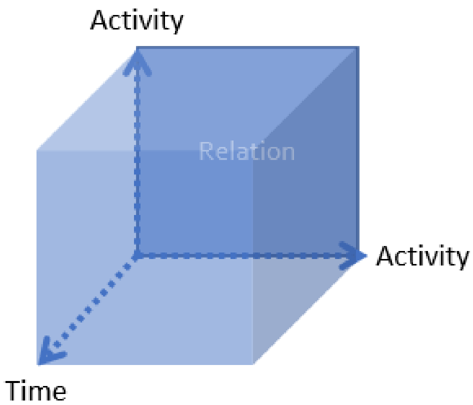

Figure 2

In practice, it is challenging to record such observations very time-accurate. Some observations make it very difficult to record such an exact time. For example, when observing emotions, the beginning and the end of a certain emotion is often difficult to determine, even if they are operationalized for the purpose of observation because different people express emotions in nuances differently. Also, the detection of very fast occurring activities is challenging, for example, in the observation of intensively playing children or

[5] For example, during the observation of surgical teams when handing over medical instruments.
[6] E.g., for facial expression analysis.
[7] Of course, only if all other observation parameters are correct. The exact timing of uninteresting or not target-oriented content is nonsensical, but an often-seen mistake.

the content analysis of dialogues in fast languages (e.g., Japanese, Spanish) (Goethe 2014). Ultimately, "observing" can be expressed by the following definition:

Definition 3: Observing finally means *"to measure something in relation to time"*.

As already described, measuring here has a certain degree of uncertainty inherent in the observation method.

4 Performing an Observation

In the following, it is clarified with which goal, or more simply expressed "why" something specific is to be observed, which "something" is to be measured (the "what") and finally, "how" the observation is to be performed.

4.1 Aim of the Observation - the "Why"

Before starting an observation, it is fundamentally important to think about the goal of the observation. That "why" such a study should be conducted at all. This may sound banal, but the lack of this answer is a common mistake in practice. An approximate goal or a supposedly known goal offers much scope for inefficient and ineffective work. As an aid, the following types of research studies can be used as a starting point for further project design:

Descriptive or comparative study: Its aim is to systematically describe the study's subject, either to collect data for later hypothesis development or to compare the collected data with data from other studies. Here, the "what" and "how" of the observation are known.

Hypothesis-Driven Study: Its goal is to support or refute the established hypothesis with systematic evidence. The "what" and the "how" are partially known here since the hypothesis rarely comes from nowhere but was developed based on known studies.

Exploratory study: Specific content or phenomena are investigated with the aim of obtaining data for planning further descriptive or hypothesis-driven studies. The "what" and the "how" are largely unknown since they are novel, previously unexplored contents/phenomena.

The goal must always be to determine the "what" and "how" of the observation as accurately as possible for all types of studies because they are important for

- the duration of the observation,
- the number of results,
- the quality of the results and
- the meaningfulness of the results.

Thus, the two variables "what" and "how" provide the leverage for effective and efficient work in observational studies!

4.2 The Contents - The "What"

What is to be observed can usually be represented as a system of concepts (or terms), whereby such concepts represent the contents of the observation (e.g., "take", "show", "positive"). One often speaks of a category, coding, or rating system. Groups of concepts can be combined if they are logically related. For easier understanding, we call the observation contents"codes" and the groups of certain codes "classes". Such classes can be descriptive (a possible notation for a class with codes could be: Emotion [cheerful, reserved, fearful]). However, they can also represent scales (e.g., Emotion scale [1..5]). It is not recommended to define numerical representations, acronyms, or any other abbreviations, because in practice, a transfer is always necessary to find out what, for example, Emotion [3] means. A descriptive coding system with terms such as Emotion [smile] is easy to communicate, easy to learn by new project members, and largely unambiguously applicable. This is a major advantage because, already in the early phase of data acquisition, the quality of the data (few misinterpretations/misunderstandings) and the entire observation process's speed can be influenced positively. Of course, both forms of code definition require the most exact and detailed description of each individual code. E.g., *„Emotion [beginning of a visible smile] is given if the person shows an obvious reaction of joy, for example, an upward-moving corner of the mouth (on one or both sides) or forming laughter lines at the eyes. Other characteristics are..."*.

It is very useful to define the codes of a class according to the following criteria:

1. Atomic: The code describes exactly one fact or phenomenon of reality. It represents information that cannot be further subdivided. The code "positive verbal expression", for example, would violate this criterion. Because in order to evaluate later how often the person has made (any) verbal utterances and how long this person has been in a (any) positive mood, all such recorded codes would have to be broken down into their parts later in a time-consuming process. The assignment of Mood [Positive] and Expression [Nonverbal], on the other hand, is precise and simple.

2. Mutually exclusive: Only one code from a class can occur in real life at a time, e.g., in the class Position [sit, lie, stand, walk], because an individual cannot sit and walk at the same time. Otherwise, there will be overlaps later in statistical evaluations causing problems in the data's meaningfulness.

A coding system defined in this way can contain any number of classes, each with any number of codes, and can be structured hierarchically or even recursively. This makes it possible to map any complex facts or phenomena in detail in the form of a coding system. In this context, a software tool such as Mangold INTERACT (INTERACT 2017) is of great benefit, as it greatly simplifies the administration and application of such coding systems (see Glüer 2015; Hollingshead & Poole 2012; Mangold 2000, 2004, 2005, 2006; Mogge 2008; Allen, Lehmann-Willenbrock, Rogelberg 2015).

The definition of such systems resembles a language in which letters are transformed into words, words into sentences, and sentences into paragraphs. That is why INTERACT speaks of *lexical chains*, which make a complex subject easily describable. Figure 3 shows an example of such a coding system.

The use of descriptive terms is advantageous in practice. The user can easily memorize the sequence of codes, as they are terms of their common language. Such a procedure speeds up the data entry process and makes it robust against incorrect data collection.

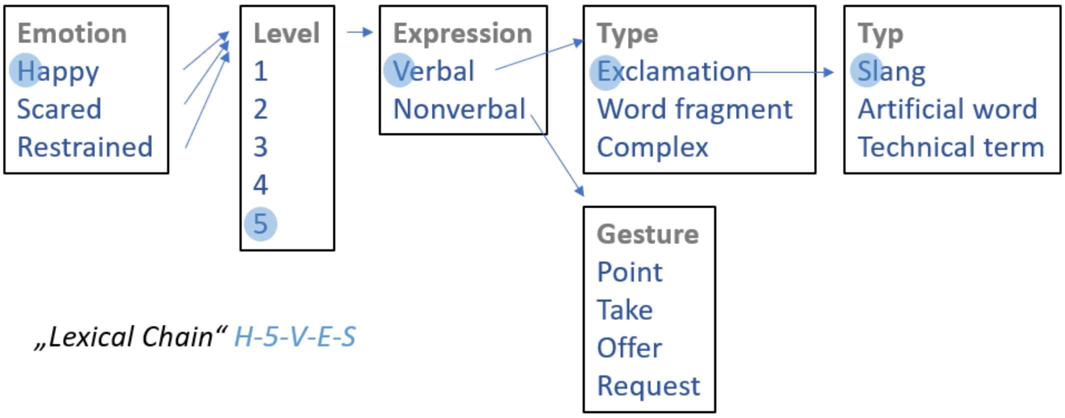

Figure 3: Coding System and Lexical Chain

For the sake of completeness, the recursive definition should be mentioned, where a few classes are linked in a circle (e.g., made possible by further link-information in INTERACT). This way, complex codes can be assembled during input without having to be defined before. For example, if the Facial Action Coding System (FACS) is used to describe facial expressions (Ekman 1978), the coding system would have to consist of well over 100 codes to be able to code each muscle, its direction, and intensity of movement. By recursive linking, this can be reduced to three simple classes (Action Unit Digit [0..9], Intensity [a,b,c,d,e], Laterality [right, left, symmetric, asymmetric, top, bottom]). This is not only an enormous facilitation but also a practical necessity because, in software-supported data acquisition, the computer keyboard is the limit for associations of keys to codes. Depending on the national language, these are approximately 26 - 29 letter keys and 10 number keys.

When defining a coding system, it is crucial that the information necessary to answer the research question is really captured. This is far from trivial because the same situation can be observed in many different ways. If, for example, team discussions (see Schulte, Lehmann-Willenbrock, & Kauffeld 2015; Schneider, Liskin, Paulsen, & Kauffeld 2015; Lehmann-Willenbrock, Chiu, Lei, & Kauffeld 2017) focus on the content of a conversation, there may be a lack of data later as to which persons had which emotional mood on which topics. If, on the other hand, attention is paid to the speech behavior (inviting, questioning, defensive, etc.) of the individual team members (project manager, team leader, CFO, etc.), important data may be missing at a later stage in order to be able to answer which topics cause the mood to change emotionally. Therefore, it is indispensable to think about the design of the coding system in advance because the option to capture “everything possible” to be able to examine all possible data combinations later is practically impossible and the wrong use of resources[8]. It is advisable to limit oneself to a few “atomic” codes because these represent one single fact each, they are unambiguous, and thus can be collected very quickly[9]. By combining these codes, questions can be answered that would be very difficult or impossible to capture by observation. If, for example, the question is to be answered which persons look at which objects and under which verbal content (e.g., in a school environment or in development studies with children and their parents), the data collection of verbal utterances and individual gaze direction is possible quickly and independently of each other. By using a software-supported data acquisition and analysis tool, a simple combination of data can be made with just a few mouse clicks to determine which objects all persons are looking at simultaneously when these objects are verbalized at the same time. If one wanted to observe all periods in which object X is spoken about and each of the persons' present looks at the object, the observers would have to constantly coordinate their attention between hearing and seeing and focus on different visual areas. This is hardly feasible in practice[10].

4.3 The Procedure - The “How”

The above remarks suggest that a live observation of the event to study is probably very difficult, if not impossible. Especially the observation of several atomic codes requires the multiple iterative observation of the same real-life situation. Also, observing the smallest phenomena (e.g., eye-rolling, eye contacts, subliminal touches) or complex processes (simultaneous speaking and acting) require a recording of the situation to be observed. This is best done audio-visually using the latest video and audio technology available on the market. This topic cannot be dealt with in any further depth here, as it is too complex. However, it is recommended to seek advice from a company that specializes in behavioral observational studies! With this topic, there is the danger of putting a heavy strain on the financial budget, time, and ultimately the nerves of all those involved, without receiving a corresponding benefit.

The advice of a typical audio-video system provider or the AV department of a research institute is only recommended to a limited extent. They are usually very familiar with the latest video technology, but in some cases, not with the needs of complex research studies. For example, there are many video laboratories in which cameras are mounted on the ceilings. If the research focuses on early childhood development (e.g., Striano 2016), such positioning often makes little sense. Apart from recording the participants' heads from above, little can be seen in such video recordings - certainly not the desired facial expressions of the test subjects. From almost 30 years of experience it is very clear that demands for scientific

[8] Much too time-consuming and error-prone due to physical and emotional fatigue of project participants.
[9] Observers can easily be trained to grasp simple facts and phenomena, so that they can finally grasp them almost mechanically and with high accuracy.
[10] This misapplied methodology is one of the reasons for the still negative reputation of observational studies. Because those who have had to do something like this before, know how frustrating and time- consuming this type of data collection can be.

observation are unequal to demands of surveillance. Even the purchase of state-of-the-art technology is often unnecessary and a cost trap. For example, having all cameras remotely controllable is a nice thought. In practice, however, the question arises as to what use the control system is in a room that can be seen completely in the picture anyway, and who is to carry out this control during the experiments. It is therefore strongly recommended to consult a specialist beforehand. It is also advisable to obtain an opinion from colleagues with similar research projects. This ensures that all practical problems that have already arisen there can be discussed in advance with the professional provider. The latter probably already has newer / better solutions since the lifetime of audio-video and computer technology is typically very short.

Suppose the observation situation has been recorded with several cameras and microphones from different angles. In that case, the question arises how these can be evaluated as quickly and purposefully as possible. A professional and specialized software tool is indispensable here. The usual computer software in the video field and other common software tools such as SPSS, R, or office applications are not designed for such a task. A tool that can perform this task is, e.g., the software INTERACT (Figure 4). In use for decades (since 1991) in various application domains worldwide (Mangold 2017), this program contains a wealth of special routines specialized in acquiring and analyzing observation data. Also, the integration of data, such as EEG or physiological values, is possible. Thus, qualitative and quantitative data can be combined, which allows an easier and deeper understanding of the individual, previously independent data streams.

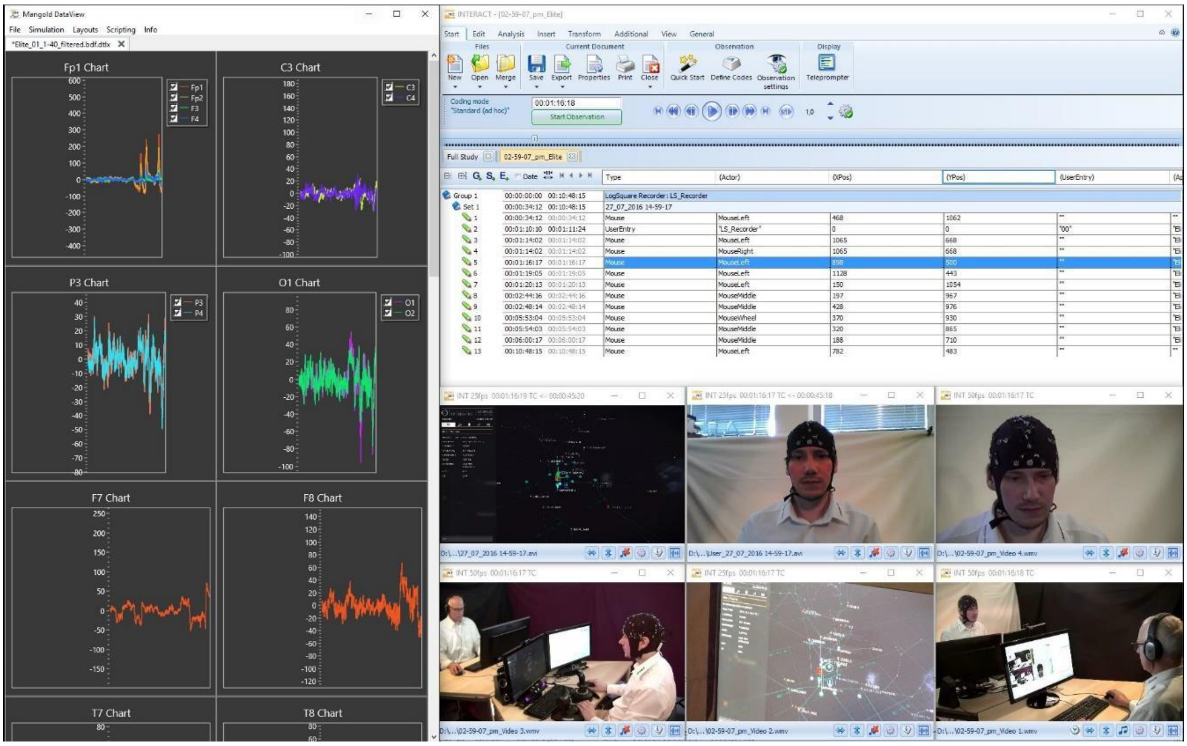

Figure 4: Mangold INTERACT®, with DataView add-on module

The following procedures of observation and data collection are common:

The so-called time sampling (Figure 5): How it works: The time axis of the videos are divided into equidistant intervals. Then the end (or the beginning) of each interval is jumped to, and the still image at this point is described with the previously defined codes. This method is traditionally used in ethology, where little change can be seen over long periods of time (e.g., animals in stables). In this case, it is not very useful to watch hours of videos (sometimes recordings over several days!) of lying or standing animals if the desired information can be recorded in 5-minute steps. However, time sampling is also often used by playing such intervals (several times in a row) and then performing a global evaluation of the interval using the codes. This procedure seems simple at first but may be born out of an inaccurately or too complexly defined coding system, which makes it difficult to define concrete points in time for the start and end of a code. More serious, however, is the fact that valuable information is lost in the process, and incorrect information may be created in the data. We refer to such false information as ghost information since it usually only becomes visible by combining existing data (see Figure 5 and Figure 6[11]). It should be noted that this form of data acquisition results in incorrect frequencies of codes. If a series of intervals is coded with an identical code, that code's overall duration may be correct, but not its frequency of occurrence. The code only appeared once in reality but was broken up into several fragments by the time intervals, which may lead to misinterpretation. Such a global evaluation also contradicts the scientific objectivity discussed at the beginning of this paper, which is intended to depict the event as precisely as possible and neither interpret nor generalize it. To use time sampling also for reasons of better rater reliability is questionable[12] for many reasons. Nevertheless, there may be research topics where such a time sampling approach can be useful. However, the event sampling method is preferable because it more accurately maps reality into data.

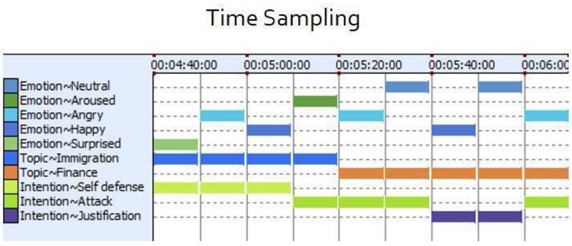

Figure 5: Time Sampling, INTERACT Time-Line Chart

The so-called event sampling method: Here, observed facts or phenomena are recorded with their exact start and end times. This method appears to be more complicated but is very easy if an appropriately simple coding system is used. Several persons can independently collect atomic observation

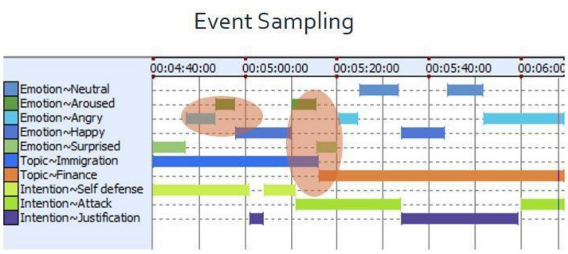

Figure 6: Event Sampling, INTERACT Time-Line Chart

[11] Figure 6: From a campaign speech, during the "immigration" theme the emotions Angry, Aroused and Happy appeared in real life, while in the time sampling variant (Figure 5) only Angry is coded, as the predominant emotion in this interval.

[12] In particular, the arbitrarily definable interval width decides on the quality of the results. INTERACT e.g., evaluates the rater reliability independent of intervals, based on real code frequencies and temporal overlaps.

data, which in practice can be done very quickly with the right software support and some tricks[13]. An essential advantage is the possibility to distribute the data collection to any number of persons. The procedure can be standardized by the software used in the working group. At the same time, it is possible to easily copy the videos in the form of files so that all project participants can work on the same material. This way, specialists for certain coding criteria can be trained in the group, who then work on the videos in parallel. The software ensures that all data can be put into a large context at the end of the data collection process by simply merging them. The complex data thus obtained can now be evaluated with software support to gain new insights. The software INTERACT also offers the possibility to export event sampling data as time sampling data. So other programs that need time sampling data can also benefit from this much more precise data. In particular, the interval width over the data can be set arbitrarily for different research needs, leading to different results. This way the research group can get the best out of both worlds by using a single coding method.

The *automatic coding* (Figure 7): This method is useful for certain observations, but the method's robustness must be checked in advance. Otherwise, there is always doubt about the reliability and validity of such automatically collected data.

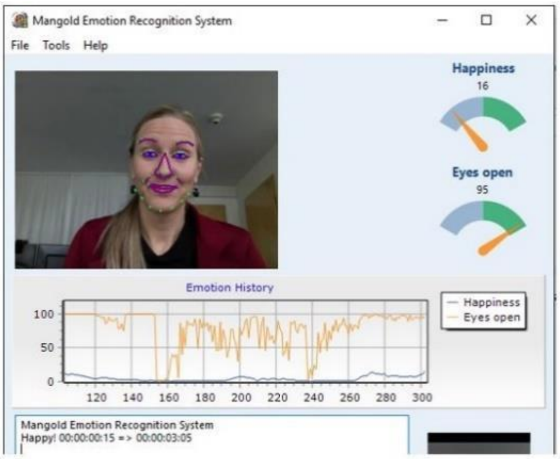

Figure 7: Automatic Coding, Mangold Facealyzer®

Transcription (Figure 8): Transcribing video or audio files is a widely used method. The main advantage is being able to analyze the recorded texts from a linguistic or statistical point of view. The use of transcription to describe the observation from the observers' point of view may be of limited use. Because to generate useful statements from the transcripts, sooner or later, they must be coded schematically in some form on the way to knowledge acquisition. In such cases, event sampling instead of transcription is a good choice. This saves a lot of time and probably provides more information since the necessary coding is already done during the observation and not afterward, based on the transcripts.

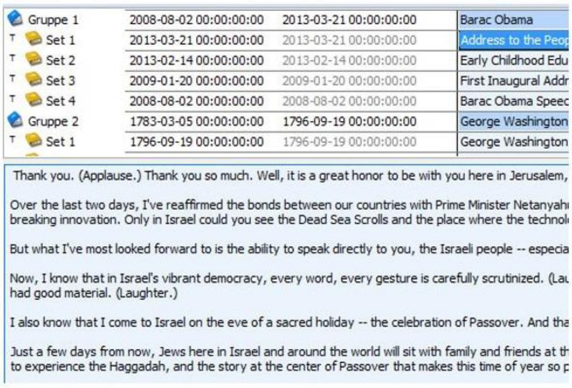

Figure 8: Transcription, INTERACT

[13] For example, video segments can be fast-forwarded if the object to be observed is not in the field of view. Videos can also be played faster, whereby the speed is only reduced at the necessary points. Also, a well-considered use of keyboard shortcuts for data input and video control can simplify the process considerably.

5 Methods of gaining knowledge

If data was collected as described, it can now be converted into meaningful results. In a relatively complex process, the observed reality was mapped into codes (Figure 9), which are now converted into descriptive statistics, e.g., the frequency, duration, and the percentage of codes over the observations ("how often does A laugh, how often does B laugh, how long does A touch, how long does B touch" or similar). This is certainly useful. However, the use of such data is limited.

Figure 9: Schematic representation of a coding, over time to the right

For a fundamental gain in knowledge, a specialized software tool is necessary. One that allows a *deep insight* into the collected data and thus the observed situation. The aim is to *discover* things that could be observed only under great effort or not at all. These are usually co-occurrences and sequences of codes and specific time periods resulting from observational content. A question could be, e.g., "How long does it take for A to laugh and B to touch until B laughs and A touches" (see the vertical bars in Figure 10). At the same time, the question of contingent actions/reactions that occur in a defined time interval as a consequence of a preceding (complex) action arises. Furthermore, it is necessary to define when these contingencies are to apply, from the beginning of the successor (e.g., one person interrupts the other while speaking) or until the end of the successor (e.g., a question has been answered completely).

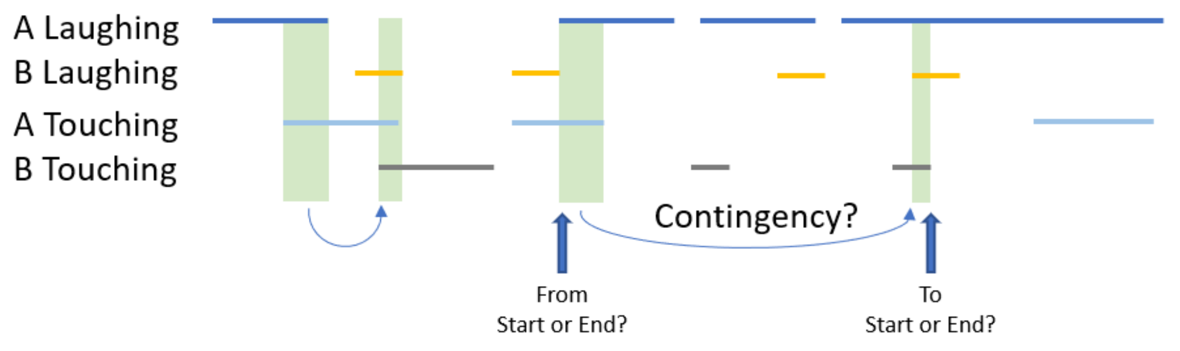

Figure 10: Highlighting of simultaneous codes and contingent sequences of such code combinations

It can also be interesting to investigate only those periods that cross such contingency intervals (Figure 11). One question could be: "What happens during these contingency actions?" For this purpose, the codes that are precisely trimmed to contingency intervals can be automatically turned into statistical values. They can also be converted into new codes, which now have a much more complex meaning and therefore could not be coded during observation.

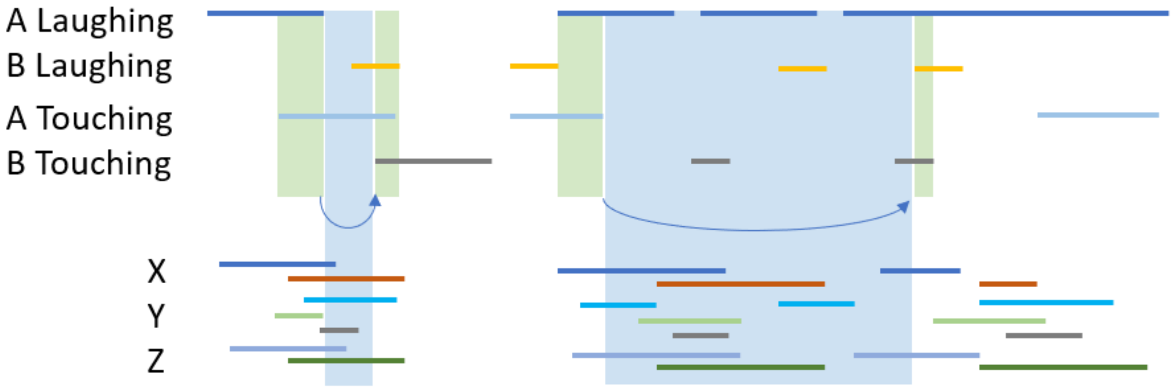

Figure 11: Highlighting of the found contingency intervals

This arbitrary repeatable sequence of evaluation steps (search for simultaneities, sequences, pruning of codes to intervals, generation of new codes based on these evaluations) offers almost unlimited possibilities to generate arbitrarily complex answers from originally simple, atomic data. Thus, facts can be discovered that were simply invisible in the original observation due to the observer's naturally limited cognition. Not only the comfortable acquisition of the raw observation data, *but more importantly, this possibility for information mining is the real added value of such specialized observation software*.

It is important to mention that such evaluations can typically only be carried out in a very complicated way with other common tools (e.g., Matlab, R, SPSS, Excel), or in some cases not at all. After exporting the raw codes into such a program, the exact time correlation is usually lost. Also, some software tools cannot process the data on a time axis because they are case-oriented. This often means that the entire added value is lost because the data was exported too early due to traditional tool-use. The well-known statistics programs, however, have their uses since they offer procedures that a specialized observation tool does not provide. However, they should only be used if the possibilities of the observation tool have already been exhausted.

6 Summary and discussion

As shown, answers to complex research questions can be generated from simple observation data if certain procedures are followed, and conditions regarding content are observed. The use of a software tool specialized in observational studies is useful in every project phase. Correctly applied, it generates an added value that cannot be generated in this way by any other, not specialized software.

Further topics to be investigated are especially the connection between semantics and practical application of codes and the related statistical evaluations. Especially the observer agreement is strongly affected by this. For it is open, which meaning in the sense of reliability, e.g., the assignment of a code “dog barks” has, if different observers can assign it in any frequency (Observer A codes “dog barks” as one long event, while in extreme cases observer B codes each single short barking sound as individual event). The comparison of both data sets leads to completely different statements depending on the method. Furthermore, it has to be investigated how the semantics of arbitrary codes, which can be freely defined by the user, can be represented in a way that is understandable for the computer so that the above-described analyses can be performed automatically or semi-automatically.

Sources

Allen, Joseph A., Lehmann-Willenbrock, Nale, Rogelberg, Steven G. (2015). The Cambridge Handbook of Meeting Science. Cambridge University Press. ISBN 978-1-10758-973-5

Ekman, Paul., Friesen, Wallace. V. (1978). Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto, California: Consulting Psychologists Press

Faßnacht, Gerhard (2007). Systematische Verhaltensbeobachtung. Stuttgart: UTB Verlag

FH Münster (2017). Institut für Berufliche Lehrerbildung. https://www.fh-muenster.de/ibl/projekte/informationsportal/Beobachtung.php, Last check 2017 May 01

Gabler Wirtschaftslexikon (2017). http://wirtschaftslexikon.gabler.de/Definition/prozess.html, Last check 2017 May 01

Glüer, M. Software zur Videodatenanalyse: Ein Überblick. http://www.videodatenanalyse.de/software, Last check 2017 May 01

Goethe Verlag GmbH (2014). www.10winds.com/50languages/did_you_know/DE042.HTM, Last check 2017 May 01

Greve, W., Wentura, D. (1997). Wissenschaftliche Beobachtung. Eine Einführung (2., korrigierte Auflage). Weinheim: Psychologie Verlags Union

Grossmann, Paul (2004). Das Üben von Achtsamkeit: Eine einzigartige klinische Intervention für die Verhaltenswissenschaften. In Heidenreich, T & Michalak J. Achtsamkeit und Akzeptanz in der Psychotherapie. Tübingen: dgvt-verlag. Kap 6.

Hollingshead, Andrea B., Poole, Scott (2012). Research Methods for Studying Groups and Teams. Abingdon, Oxon, UK: Routledge (Taylor & Francis)

INTERACT (2017). Software Description. https://www.mangold-international.com/en/products/software/behavior-research-with-mangold-interact, Last check 2017 May 01

Kabat-Zinn, J. (1990). Full catastrophe living: Using the wisdom of your mind to face stress, pain, and illness. New York: Dell.

Klonek, Florian, Kauffeld, Simone (2016). Watch Your Language! Analyzing active ingredients of client speech in a Motivational Interviewing Intervention for Environmental Behavior Change. Technische Universität Braunschweig, Institute of Psychology, Braunschweig

Lehmann-Willenbrock, N., Chiu, M. M., Lei, Z., & Kauffeld, S. (2017). Understanding positivity within dynamic team interactions: A statistical discourse Analysis. Group &

Organization Management, 42, 39–78. doi:10.1177/1059601116628720

Mangold P. (2000). Softwareunterstützte Verhaltensbeobachtung in klinischer Arbeit und Forschung. A. Horsch, Bildverarbeitung für die Medizin, (pg 444-448). listed in dblp, Uni-Trier, ISBN 1431-472X, Informatik aktuell, Springer Verlag.

Mangold P. (2004). Speed up research process, getting better results and reducing costs at the same time can go hand in hand very well. Advances in Ethology 38, Contributions to the 5th International Symposium on Physiology, Behaviour and Conservation of Wildlife, Berlin, Germany, (pg. 64), Blackwell Verlag Berlin.

Mangold P. (2005). Proceeding studies on behavior - not only a challenge for professional tools. Empowering Software Quality: How can Usability Engineering reach these goals? 1st Usability Symposium, HCI&UE Workgroup, Vienna, Austria. books@ocg.at 198 Austrian Computer Society, ISBN 3-85403-198-X (pg 127-140), listed in dblp, Uni-Trier

Mangold P. (2006). Softwareunterstützte Audio- / Videoanalyse und Multimodale Datenauswertung in Beobachtungsstudien. Usability Professionals (pg. 109-111), German Chapter of Usability Professionals Association e.V., ISBN-10:3-8167-7177-7

Mangold P. (2006). Getting better results in less time. When using audio/video recordings in research applications make sense. Abstracts of European Society of Family Relations, 3rd International Congress (pg. 73), Voss H. (Etd.), Darmstadt, Germany

Mogge, Sabine (2008). Videoanalyse von Diskussionsprozessen und -inhalten zwischen Grundschülern. Kassel, University Press, ISBN 978-3-89958-432-5

Norris, S.P. (1984). Defining observational competence. Science Education, 68(2), 129-142, John Wiley & Sons, Inc.

Schneider, K., Liskin, O., Paulsen, H., Kauffeld, S. (2015). Media, Mood, and Meetings: Related to Project Success? ACM Transactions on Computing Education, 15(4), 21.

Schulte, E.-M., Lehmann-Willenbrock, N., Kauffeld, S. (2015). Treat us fairly and we won’t complain: Multilevel effects of procedural justice on team meeting interaction. Psychology, 6, 1795-1810. doi:10.4236/psych.2015.614176

Striano, T. (2016). Doing Developmental Research: A Practical Guide. Ney York: Guilford Press

Further reading

Arias, Anna Maria; Davis, Elizabeth A. (2016). Recording Observations: When Done Well, Observations Can Serve as Evidence When Engaging Science and Children. v53 n8 p54-60 Apr 2016 in Science Practices.

Curby, Timothy W.; Johnson, Price; Mashburn, Andrew J.; Carlis, Lydia. (2016) Live versus Video Observations: Comparing the Reliability and Validity of Two Methods of Assessing Classroom Quality. Journal of Psychoeducational Assessment, v34 n8 p765-781 Dec 2016

Cuthrell, Kristen; Steadman, Sharilyn C.; Stapleton, Joy; Hodge, Elizabeth (2016). Developing Expertise: Using Video to Hone Teacher Candidates’ Classroom Observation Skills. New Educator, v12 n1 p5-27 2016

Hussy, W., Schreier, M., Echterhoff, G. (2010). Forschungsmethoden in Psychologie und Sozialwissenschaften. Berlin: Springer

Kripalani, Lakshmi A. (1987). Observation. NAMTA Journal, v41 n3 p229-247 Sum 2016 (Reprinted with permission from “AMI Communications” 1 (1987): 2-11)

Lontz, Karey M. (2015). Observation: A Practice That Must Be Practiced. Paper presented at the “Observation: The Key to Unlocking the Child’s Potential,” North American Montessori Teachers’ Association (NAMTA) Conference, Part 1 (Denver, CO, November 5-8, 2015).

Martin, E., Wawrinowski, U. (2006). Beobachtungslehre. Theorie und Praxis reflektierter Beobachtung und Beurteilung. Weinheim: Beltz Juventa

McLeod, S. A. (2015). Observation Methods. www.simplypsychology.org/observation.html, Last check 2017 May 01

Williams, Gwendolyn M.; Case, Rod E. (2015). Tale of the Tape: International Teaching Assistant Noticing During Videotaped Classroom Observations. ISSN: 2162-3104 Print/ ISSN: 2166-3750, Journal of International Students, v5 n4 p434-446