RRREaT – Data-Analytical Techniques,
2024, 1: 01

Situations and the Search for Conditional Information

Joke H. van Velzen
Research Institute of Child Development and Education, University of Amsterdam,
The Netherlands

Abstract

In the latest RRREaT-PT editorial, and also in some other RRREaT-PT editorials and articles, the study of situation as a variable in terms of the influence of situation on objects, phenomena, and processes as it is actually happening in the world, is of interest for scientific research because it can provide for conditional information. Conditional information can inform about situational differences that matter for understanding persons, phenomena, and processes, such as the identifying of inter-individual groups based on person-specific intra-individual data. However, the precise meaning and data analysis for detecting conditional information in situations via empirical research is a complex research subject because so many variables, more than those that are being observed intentionally, can have an influence on the results. In the research literature, the concept of conditional information is mostly substantiated via theorizing rather than via empirical studies and mainly in probability theory, system analysis or network-model theory, and social-dynamics theory. In this article, I will discuss this research literature on conditional information in search of a clear definition for the purpose of data analysis in social and behavioral research that involves the variable of situation. The emphasis will lie on initial all-including data-analytical techniques to which end the four-quadrant relationship display is being proposed, with the aim to regard its usefulness for detecting conditional information in situations as it meaningful to and influences groups of persons.

1. Conditional Information: An Introduction and Preliminary Description

In RRREaT-PT Editorial 2022, Van Velzen wrote that essential issues regarding research and the continually changing actual world are: we cannot (a) know accurately the meaning of behavior, events, and trajectories if we cannot include most of the near-present influential information about what is happening in the actual world via variables and characteristic features and (b) understand comprehensively the objects, phenomena, and processes in the actual world if we cannot account for their nature or functioning, relativity of the observer, and the passing of time. These two issues are not only interchangeable, but also essential because many things are taking place in the actual world at the same time and across different timescales, and they are observed by humans from multiple perspectives, in that it becomes difficult to find meaningful patterns that provide for accurate information about the way things are functioning as they do in actual-world situations. Already, scientific research has provided for important general and detailed facts and understandings about the actual world, so that a next and worthwhile research question has arisen, namely how to elaborate towards a scientific understanding of the continuity, simultaneity, and interconnectedness in actual world objects, phenomena, and processes by including situations, and their events and states, regarding the when and why of occurring patterns.

When the study of the actual world involves situations, then both the finding of meaningless patterns and the overlooking of meaningful patterns can occur when the researcher has to account for the continuity, simultaneity, and interconnectedness that take place in the actual world. For example, patients who have food allergies, intolerances, and other food-related adverse reactions (see Lessens et al., 2020; Malone & Daley, 2024, for overviews) can experience insufferable to annoying health conditions, such as throwing up, eczema, skin rashes, gastrointestinal disorders, breathing problems, and headaches. To discover which food produces the health issues, allergy specialists can employ skin-prick tests, blood tests, food challenging (e.g., the oral consumption of varying strict doses of egg), and the elimination diet as medical diagnostic instruments. For example, the elimination diet requests patients to stop eating certain food, such as gluten, for a certain period of time (i.e., often four to six weeks), after which the food is slowly reintroduced to patients' daily eating patterns to enable an evaluation. To this end, the planning phase of elimination diets requires a thorough analysis of the dietary history, medical symptoms, past medical history, family medical history, and diverse laboratory testing (i.e., the immune-mediated IgE allergies, as well as lactose and fructose intolerances, can be assessed accurately via tests). Despite these effective diagnostic instruments for food allergies and adverse reactions, there are important diagnostic limitations due to there being different kinds of food and food components and combinations that can be found in someone's eating pattern (e.g., fats, carbohydrates, and proteins; vitamins and minerals; natural versus processed food; herbs, spices, and additives), but also the strict recommendations of eliminating certain foods in someone's eating pattern, and the emerging of nutritional deficiencies can be problematic. When there are various underlying causes for food allergy and adverse reactions (e.g., reacting to chemical components and cross contaminations involving grass and tree pollen), it can become even more difficult to entangle the different food adverse patterns via elimination diets. When also the human factor is included, such as having a preference for certain kinds of food and not being ready to acknowledge the food allergic reactions, the analysis can become obscured. Therefore, the difficulty of observing the complexity and dynamics of the actual world raises the question of how to facilitate research and, especially, data analysis with regard to finding existing and meaningful patterns in situations.

When researchers set up research studies, they can reckon with there being different kinds of values for certain kinds of variables. For instance, the value of some variables cannot vary in place and time because their value is empirically determined. Examples are the speed of light in vacuum ($c = 299792458$ m/s^{-1}) and Planck's constant or the amount of energy a photon can carry based on the frequency of its electromagnetic wave ($h = 6.62607015$ x 10^{-34} J/Hz). These kinds of variables with established values are called constants, and they mainly occur in the sciences. Second, the value of other variables is set at a certain value at the beginning of certain research. For example, velocity as a function of time can be calculated by giving a specific value to time, such as $t = 0$. These kinds of parameters that have or are given a starting value are called initial conditions. In the Merriam-Webster dictionary, initial conditions are defined as any of a set of starting-point values belonging to or imposed upon the variables in an equation. As such, an example for the social and behavioral sciences is the imposing of a specific age group for the participants at the beginning of a research study. Third, the value of the variable that is conditional information refers to additionally relevant information that consists of a stipulation or provision in terms of including the word of *if*: "This will happen *if* . . ., because . . ."

Conditional information defined as additionally relevant information because it provides for a stipulation refers to any *additional* information that has an influence or impact on the behavior, events, and trajectories of objects, phenomena, and processes. For example, Brownian motion is the random movement of colloids (e.g., grains of pollen, inorganic clay dust, particles, molecules, and bacteria with a diameter of 1.0 μm) of their own accord in fluid and gas (see Bechinger et al., 2016; Libchaber, 2019; Philipse, 2018, for overviews). The colloids react to each other and everything else in the fluid and gas until they become diffused, which requires time. The dynamics of the motion of the colloids involves (a) the shape and unevenness of colloids (e.g., the degree of being spherical, fast, and large, and their composition, such as charge), (b) the kinetic energy of colloids, (c) added energy (e.g., sunlight and stirring the liquid), (d) the collisions of colloids with other colloids and substances in the liquid and gas, (e) the ability of colloids and other objects in the liquid and gas to coagulate, stick together, and react chemically, and (f) the context or surrounding influences (e.g., viscosity or density, pressure, flow, and energy or temperature of the surrounding liquid and gas). Now, these six kinds of stipulations regarding Brownian motion have become known, however, they were not when Brownian motion was first discovered and their precise interactions are still being researched (Libchaber, 2019). Therefore, these six kinds of stipulations provide for conditional information regarding the Brownian motion of colloids because, when these stipulations are not known and not included in research studies, then only a general diffusion equation (Einstein, 1905) can be employed to state the overall fluctuation-dissipation relationship rather than understanding when and why a diffusion deviates from the general diffusion equation as a consequence of what is actually happening in certain situations.

Conditional information that is not yet known refers to information that needs to be sought in addition to, and that can arise unexpectedly regarding, the expressly gathered data. Conditional information always encompasses relevant information for data analysis because it provides for meaningful descriptions and explanations (see Van Velzen, 2024a). From a data-analytical perspective, not yet known conditional information may reside in (a) latent variables (i.e., the underlying properties and mechanisms statistically inferred from the observed variables) and emerging variables (i.e., the uncovering of something new that was already there, but it was not known and not present as an underlying property of the observed variables) and (b) partial, spurious, and nonlinear correlations, and it can require (c) causal conditioning procedures (see also Section 2).

This theoretical article begins with a research overview of the major theories regarding conditional information. In the research literature, there are three relatively recent and diverse theories regarding conditional information, namely probability theory, system analysis theory, and social dynamics theory. The aim of this article is to present the possibilities and challenges of detecting conditional information in research studies, which is examined from a theoretical data-analytical perspective because conditional information can clarify influences on findings and results if it is worked out appropriately. Alongside this theoretical article, there is also an empirical preliminary illustrative example on detecting conditional information (Van Velzen, 2024b). The focus of this theoretical article is on explicating precisely the way in which conditional information can clarify meanings and consequences regarding situational complexity and dynamics for the social and behavioral sciences.

2. Theories, Data-Analytical Techniques, and Conditional Information

To detect conditional information, data analysis can be employed to identify certain parts of a whole in order to interpret the data-analytical results and findings in terms of evaluating their meaning (e.g., when and why the additionally relevant information has an influence or function regarding a situational affair) to obtain knowledge (i.e., if . . ., then . . ., because . . .). Most actual-world data can feature complex dependencies between variables and include continuous, categorical, as well as unobserved latent and emerging variables. Examples of recent data-analytical procedures for obtaining conditional information are (a) conditional clustering (i.e., partitioning data into interpretable subgroups: Molnar et al., 2023), (b) leaving out one covariate to evaluate the changes in results (Lei et al., 2018), (c) local and global explaining of the role of the variables in the results (Covert et al., 2020), and (d) sequential knockoff variable selection for mixed data (Blesch et al., 2023).

To detect conditional information, in that it contributes to the meaningfulness of certain observed patterns in the data, a first self-evident data-analytical technique is the determination of relationships via covariance, correlation, and regression calculations between two or more variables. Although these data-analytical techniques provide for information, each has a serious limitation, namely covariance does not take into account the size of the variance of the variables and, thereby, only informs about how they vary together (i.e., deviate from the mean score) and both correlation and regression do not inform about causality (i.e., if and how the one affects the other). For example, a high correlation between hair color and musical skill does not provide for a reasonable empirically justifiable connection. Similarly, the regression line merely represents the average relationship for a whole group, which means that most individuals will not exactly fit the regression line. Correlation coefficients and regression lines can be prone to: (a) providing for spurious relationships, such as correlations being sensitive to outliers that can substantially inflate or deflate a correlation, and the found correlations may be confounding because there is an underlying, unmeasured common cause and (b) failing to detect a relationship, such as the correlations having a nonlinear relationship (e.g., u-shaped curve), the inappropriate combining of groups that should have been considered separately in order to not mask the relationships of the individual scores (e.g., only part of the range is sampled), and the essentiality of both variables having to have some variation to produce a correlation, in that it can be just a coincidence.

To determine causality, specifically designed experiments need to be conducted that include all *if*s, similar to how Brownian motion experiments need to account for the shape of, kinetic energy of, and the collisions and interactions between colloids as well as the added energy and surrounding influences. To set up specifically designed experiments require preliminary work via prior studies. This raises the question of how to detect conditional information via preliminary research studies when the research subject is complex and dynamic. The following three sections present the major theoretical accounts regarding the detection of conditional information and the complexity and dynamics of the actual world.

2a. Conditional Information and Probability Theory

Probably, the best worked out definitions of conditional information, in terms of being theoretically substantiated and mathematically formalized, reside in probability theory (see Jaynes, 2018; Knuth, 2016, for historical overviews). Probability theory refers to the (mathematical) plausible reasoning regarding the inferences that are made based on observations from the actual world. Probability theory is grounded on the assumption that the likely occurrence of something to happen in the future agrees with the relative frequency of repeatedly envisioned observations. For example, the actual observation of the flipping of a coin for multiple times produces the relative frequency outcome of 51 percent (i.e., 0.508: Bartoš et al., 2023) for landing heads or tails. In probability theory terminology, the envisioning of flipping coins produces the probable outcome of heads or tails of 0.5 or ½, in that all probable outcomes in a situation must add to one (i.e., $0 \leq P(A) \leq 1$). Conditional probability refers to the related information of dependent events (i.e., $P(A \mid B)$ or A is true given that B is true). The emphasis of the major modern probability theories as they are arranged in formal mathematics ranges from the relative frequency that an event can occur, to degrees of knowledge about events, and to subjective agents' best guess about the occurrence of an event. Especially, these three major theories differ with regard to dealing with conditional information.

The first modern probability theory as a formal mathematical theory about the relative frequency that an event can occur is from Kolmogorov (1933) and it is also referred to as a Frequentist probability theory. Kolmogorov's theory and axioms for probability depart from analogous properties of sets, in which there is a sample space Ω of elementary events in a field F that consist of certain selected subsets. Its axioms state:

(a) Normalization or $P(\Omega) = 1$;
(b) Non-negativity or $P(f_i) \geq 0$ for all f_i in F;
(c) Additivity or the uncommon elements of F (i.e., $\{f_1 \ldots f_n)$ then $P(f) = \Sigma_i P(f_j)$;
(d) Continuity at zero or if a sequence of f tends to the empty set, then $\mathrm{P}(f_j) \rightarrow 0$.

Especially, axiom (c) regarding additivity is sometimes referred to as a functional equation that is necessary for the application of Kolmogorov's theory to distinguish differentiable meanings in the data. For example, if you know only the derivative of a function, you do not have enough information to determine it completely, but you can seek for a solution that is subject to some additional condition(s) that uniquely specify the value of the arbitrary constant(s) in the general solution.

A limitation regarding Kolmogorov's theory is called the Borel-Kolmogorov paradox, where a seemingly well-posed problem appears to have many different correct solutions, but actually involves erroneous calculations (i.e., obtaining non-sensible inferences and violating the rules of inference) by specifying, for instance, $n = 0$, which can result in jumping directly into infinite sets without considering any limiting process from a finite set. De Finetti (2017) and Jaynes (2018), among others, have argued that not only Kolmogorov, every other probability theory faces the danger of carelessly using infinite sets and quantities when it does not first establish any finite nth stage. That is, the infinite calculations can lead to cancelling out of terms erroneously those terms which are in fact necessary for obtaining the correct result. Therefore, it is not obvious whether Kolmogorov's axioms (c) and (d) are logically independent because in the set-theory context axiom (d) can provide for consistency, but in general applications, most problems can be solved by employing axioms (a) and (b) and the additivity of axiom (c).

The second modern probability theory as a formal mathematical theory about the degrees of knowledge about the occurrence of events is from Bayes (1763) and it emphasizes objective logical interpretation with regard to knowledge states. That is, where Kolmogorov's probability theory particularly refers to additivity (i.e., axiom c), Bayes theorem refers to prior probabilities or the conditional information at hand. For example, Bayes' formula is as follows.

$$P(A_i|B) = \frac{P(A_i)\, P(B|A_i)}{P(A_1)\, P(B|A_1) + P(A_2)\, P(B|A_2) + \ldots + P(A_n)\, P(B|A_n)}$$

Bayes' formula shows that the probability of A given B is true when all the conditional evidence at hand is included in the calculation. All the evidence at hand refers to all the prior probabilities that we know of. For example, Bayes (1763, p. 2) wrote in a letter to introduce this formula (i.e., according to Price, who communicated Bayes' essay), that "At first, in thinking about this subject . . . a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circumstances, it has happened a certain number of times, and failed a certain other number of times . . . it would not be very difficult to do this, provided some rule could be found . . . [that] should lie between any two named degrees of probability." A limitation regarding Bayes theorem is that when insufficient prior probabilities are available, it can become difficult to discover the range of correctness of the found probability.

The third modern probability theory as a formal mathematical theory about subjective agents' best guess about the occurrence of an event is from De Finetti (2017) and it emphasizes the subjective nature of probability. De Finetti (2017) argued that two events are always different, but they are equally probable if an individual judges them as such. In other words, probability calculation depends on the event, the person who selects subcomponents from the event, and the state of the information available regarding what is happening in the event. De Finetti (2017, pp. 22-24) added that "regarding any situation, there will always exist a number of conceivable alternatives. Depending on one's information and knowledge, some can be excluded as impossible . . . The others will remain possible for You, but neither certainly true, nor certainly false." De Finetti's theorem is as follows: If the sequence P^1, P^2, . . . is exchangeable, the P^n can be written as

$$P^n(x_1, \ldots x_n) = \int_{\Delta_d} \mathrm{d}\mu(p)\, p(x_1) \ldots p(x_{\mathrm{n}})$$

where Δ_d is the set of all probability distributions over the outcomes {1, . . . , d}, μ is a probability measure on Δ_d, μ is independent of n, and μ is unique. Hence, De Finetti's theory describes Kolmogorov's additivity of axiom d differently by including anybody's personal opinions and then search for coherence. A limitation of this definition of additivity is that a coherent probability will depend on the persons who are available to present their personal opinions.

In probability theory, the emphasis lies on logical (in-) dependence (i.e., the simultaneous tossing of two dice, and the tossing of one die, respectively), rather than on causal dependence (i.e., the cause precedes and is related to the result while no plausible alternative explanations can be found). Regarding logical (in-) dependence in probability theory and application, two interpretive ideas have been formulated: (a) frequency in an ensemble and (b) a reasonable expectation (see Cox, 1946, for an overview). Frequency in an ensemble is based on the assumption that the event is capable of indefinite repetition. For example, the ensemble of a box with three identical balls, of which two are red and one is blue, produces a 2/3 chance of drawing a red ball on a single trial, and on two successive drawings the chance on drawing another red ball is 2/3 times 1/2 is 1/3 chance. In other words, probability as frequency in an ensemble always produces calculations that are interpreted as definite results because it is assumed to be indefinitely repeatable. Probability as reasonable expectations is based on the assumption that something is likely, and the likeliness will increase when the number of instances increases. For example, ensembles or events can be artificial when our knowledge is incomplete (e.g., the tossing of a coin varying from one trial to another, and because we cannot do all trials, the probability will lie within a certain limit: Bartoš et al., 2023). Therefore, probability theory *and* application is about calculating the chances of events occurring under certain conditions, which often involves our state of knowledge, with the intention of the calculations being as much as possible without being subject to human caprice.

However, the greatest issue regarding the translation of probability theory into applications is that the basics are *not* substantiated sufficiently in agreement with the actual world. That is, where De Finetti (2017) argued that two events are always different, I would like to rephrase this argument as two events can be different in certain ways and they can also be similar in certain ways. For example, the colloids in a fluid that produce Brownian motion by moving randomly until they are diffused in the fluid, the amount of time it will take to diffuse will depend on the circumstances. Similarly, although empirical experiments may not directly reveal exactly what is happening in the actual world, well-designed laboratory experiments can produce for replicable results that inform about what is happening under which conditions. That being said, there is still the issue of translating probability theory into applications with regard to the devising of experiments. For example, the tossing of a die does not only require a fair die, but also a “fair” table on to which to throw the die, a “fair” hand or mechanical device that does the throwing of the die, and a “fair” environment with similar physical conditions regarding the throwing of the die. I have placed the word of fair in between quotation marks to indicate the issue of obtaining a fair context for the throwing of the die. Therefore, I can agree with Jaynes (2018, p. 338) who wrote that “the answer [obtained via probability] is no more than a correct conclusion from the information that was given,” in that in reality we will require all, or at least most, of the prior information that is essential to make an accurate actual-world description and explanation.

2b. Conditional Information and System Analysis

Varied scientific disciplines, such as weather forecasting, artificial intelligence, biology, healthcare, and psychology, all make use of system analysis to detect, among other things, conditional information (see Alon, 2007, for an overview). System analysis refers to data-analytical techniques that employ models to construct systems or networks of connected nodes to study their interrelationships because these systems can simplify the complexity and dynamics of trajectories, such as events in situations and behavior. That is, the model defines what is and what is not included in the system (i.e., the nodes) as well as the characteristics of the system (e.g., whether it has hierarchies and is nested), to examine how the nodes and interrelationships in the system can react over time. In this way, the function of trajectories can become better understood, but it also often remains challenging to understand the results of system analysis, especially, when multi-level systems are involved, each changing over time while unidentified conditions and indirect activation can play a role.

The detection of conditional information via system analysis, such as detecting the factors that influence the Brownian motion of colloids in fluids, requires data-analytical techniques that can inform successfully about the influencing factors. For example, the aforementioned conditional information regarding the Brownian motion of colloids in fluids has multi-level networks, each with numerous components and their own dynamic environment and mechanisms (i.e., event or process) that can change over time, and with the additional complexity of latent and emerging variables and (re-) activity. This raises the question of which data-analytical techniques can detect the conditional information that influences the patterns in trajectories as they occur in the actual world. One data-analytical technique consists of implementation strategies in an (evidence-based) intervention context (see Kim et al., 2023, for an overview). That is, the specific mechanisms through which an implementation strategy functions to achieve optimal implementation outcomes produces essential information about how, when, where, and why the implemented strategy functions in a certain context. Regarding data analysis, explicating the functionality of an implementation strategy requires that the data analyses identify and evaluate how the network, as part of the actual world, and the mechanisms, as the process of change over time, relate to one another. To this end, most system analysis either applies new empirical data or uses computer simulations (i.e., artificial data) to validate and refine the observed results. Kim et al. (2023) argued in favor of using a so-called causal loop diagram that visualizes the hypothesized variables, new variables, and non-existent variables and their relationships, as well as feedback loops and critical balancing loops.

Researchers who make use of system analysis sometimes question if system analysis for time series data can and necessarily has to inform about the causality of the obtained conditional information (Achterhold, 2021; Choi & Lee, 2024). Granger (1969) observed that although most approaches of collecting accumulated experience use the model-building procedure, it is worth investigating whether a direct procedure of estimating the components of the cross spectrum can be found.[1] To this end, Granger (1969, pp. 428-43) argued, cross spectrum data analysis requires the taking into account of several assumptions, if someone wants to test for causality:

1. The definitions of causality and feedback, will be very general in nature, where U_t denotes all the information in the universe, assuming that the future cannot escape the past (i.e., the natural flow of time), apart from the specified series of Y_t.
   a. Y causes X: $Y_t \rightarrow X_t \ = \ \sigma^2 (X | U) < \sigma^2 (X | \overline{U\text{-}Y})$
   b. Feedback occurs: $Y_t \leftrightarrow X_t \ = \ \sigma^2 (X | \bar{U}) < \sigma^2 (X | \overline{U\text{-}Y})$

   $$\sigma^2 (Y | \bar{U}) < \sigma^2 (Y | \overline{U\text{-}X})$$

2. The definitions involve only stationary stochastic processes or stationary series. In the non-stationary case, the existence of causality may alter over time. One can then talk of causality existing at this moment of time.
3. The one completely unreal aspect of the above definitions is the use of the series of U_{t1} representing *all* available information. This assumes that the large majority of the information in the universe will be quite irrelevant or will have

no causal consequence. Hence, the accordingly obtained definition of causality is now relative to set D. If set D has three series (X_t, Y_t, Z_t) rather than the expected two series (X_t, Y_t), then spurious causality could arise.

4. It is assumed that all sets of series are normally distributed and linear predictors can be used.
5. The definitions follow a purely deterministic view of the sets of series, in that a series can be predicted exactly from its past terms or cannot have any other causal influences than its own past.

However, to observe what is happening in the actual world should require multiple time series because of the various time lags, which makes it difficult to test for causality because it is not always possible to conduct active experiments. To this end, alternatives have been proposed, each with their own solutions and problems. For example, Runge (2018) examined causal network reconstruction by distinguishing the direct from the indirect conditional dependencies among multiple time series. Runge (2018) found that the practical problems of inferring causal networks based on multivariate time series data include, among other things, unobserved variables, sampling issues, determinism, stationary regime, nonlinearity, measurement error, and significance testing. That is, although we may not need to include *all* information in the universe (i.e., U_t) to establish causality, the noise in the data that comes forth from unobserved variables and et cetera can cause for a deviation from the original process, in that the actual worlds' dimensionality, diversity of variables, and nonlinearity can hinder the correct discovery of causality. Overall, recent research (see Docquier et al., 2024, for an overview) on comparing various procedures to establish causality showed that there is not a single best procedure because it depends on the characteristics of the dataset, the measurement and quantification of observations, and how the comparison is conducted (i.e., application to actual-world case studies, in the plural).

This raises the question of whether conditional information can arise from the data, in that it can be questioned if the data can reveal the properties of actual-world situations, such as capturing the underlying temporal dynamics. Regarding machine learning, Choi and Lee (2024, p. 2) phrase this question as follows: "In practice, time series data often contain missing features with arbitrary patterns, in that some features may not be observed at any [data-collection] time step" (see also Achterhold & Stückler, 2021). By inferring a latent context variable that is only observable through the modulations it causes in the dynamics of a family of systems, one can establish a context-conditional dynamics model and then employ probabilistic computations. Regarding research subjects with more complex contexts, such as biological, neural, and social (i.e., where people are connected) networks, the various interactions at different timescales complicate the estimation, description, and stability of the network structure (see Borsboom et al., 2021, for an overview). Therefore, the difficulty of detecting conditional information, especially regarding dynamic situations that can have non-observed or missing data, is that system analysis becomes suboptimal.

2c. Conditional Information and Social Dynamics Theories

The focus of social dynamics theories is on complex human social influences that involve various contexts and interactions between social agents at different timescales (see Castellano et al., 2009; Warren et al., 2024, for overviews). Human social influences are complex and dynamic because humans are not simple entities, in that they bring out detailed and construed behavior. However, although most of the complex and dynamic human social influences do exist in social groups and can be studied, theorizing (i.e., including models for system analysis) is in general too simple to describe any real situation. Consequently, although there are numerous heterogenic micro theories of human collective social dynamics (e.g., social dynamics of attitude change, trust, COVID and mental health, social norms, status, organizations and teams, and intergroup behavior), the main macro theories (e.g., opinion dynamics, crowd behavior, cultural and language dynamics, and formation of hierarchies) agree with social networks of collective interactions of social agents in the same direction (i.e., consensus), in several directions (i.e., fragmentation), and in opposing direction (i.e., bipolarization).

Although a social agent in interaction can choose a personal response, overall, interaction tends to make people more similar because the social agents commonly have shared experiences in terms of sharing history-knowledge, culture, and language. This means that, theoretically, the data analysis goes from micro or individual levels to macro or society levels to point out features, symmetries, and dimensionalities for global behavior. Research going from theory to practice can encounter the issues of (a) inferring macroscopic theory from microscopic models and (b) recognizing additionally relevant influences, such as conditional information.

Regarding the first issue of inferring macroscopic theory out of microscopic models, in most situations the qualitative properties of large-scale phenomena do not depend on the microscopic details of the process and, hence, can describe global social behavior. The next step is trying to include the simplest and most important properties of microscopic models via empirical data. However, human social influence is a complex and dynamic subject, which may complicate this next step. For example, the review of Mason et al. (2007, p. 296) provides the following conclusion: "What the majority of these [social influence] models lack, however, is a deep understanding of individual-level psychological processes . . . What our field has often failed to do is to contextualize our robust microlevel understanding of social influence processes by explicitly situating those processes in a social situation involving multiple individuals, interacting over time, linked in social networks of friendship and influence."

Regarding the second issue of detecting conditional information, the latent and emerging information are not necessarily logical and can include noise (Castellano et al., 2009). This raises the question of how to detect conditional information in line with social dynamics theories. Commonly, researchers use field studies, laboratory studies, and simulations to study the possible variants of a model, and this can lead to the detection of conditional information (e.g., additional timescales, concepts, and contextual event, such as a fire alarm). For example, Harton and Bullock (2007) found that laboratory studies about the psychological foundations of college students' group culture suggested that information that is easily passed on meets people's individual needs and goals and invokes emotions that are more likely to be communicated to others and, hence, are factors that can propagate cultural elements.

Another example involving an application of social dynamics theory is the review of Sebo et al. (2020, p. 2) on robots interacting with groups of people rather than individuals because: "Groups exhibit unique emergent properties that cannot be fully understood by merely aggregating the behavior and characteristics of individuals" and the overall understanding of robots within groups of people is limited. Generally, groups spend more time interacting with a robot than individuals do, but also can consider the robot as an out-group. Hence, groups of people can exhibit favorable (e.g., engaging with robots and explicating one's mental states) and competitive-aggressive behavior towards robots, and in turn the robot can influence human-to-human interaction in the group. Influential factors are the descriptive characteristics of the robot (i.e., the type of robot), the robot behavior (i.e., its range of non- and verbal expressions, personality, and emotion), and the interaction context (i.e., the setting and tasks and the robot's role), all of which can cause a range of various kinds of human behavior.

This truncated overview of social dynamics theories shows that research on the micro- and macro-dynamics of human social influences is thus complicated that the various and nuanced human behavior in actual situation merely are being explored and described. For example, Szabo et al. (2022) studied the micro-social dynamics of collaborative problem solving of teams in escape rooms, as a non-interventional and minimally biased social laboratory, to extract the building blocks of successful team interactions. Their results showed a range of dimensions involved in collaborative team processes, such as socio-demographic characteristics, prior relationships between team members, conversation rules and emotionally balanced communication, and gender.

Overall, the focus of social dynamics theories is primarily on modeling, and data analysis is concerned with the identification of new concepts and the validation of theories or models against empirical data. However, there is a sparse amount of empirical data. For example, Peralta et al. (2022, p. 15) noted regarding social opinion dynamics: "This limitation has led to an explosion in opinion dynamics models without a counterpart in empirical validation." Also, Castellano et al. (2009) and Warren et al. (2024) concluded in their overviews that the complexity and dynamics of human social influences has produced many theories and some varied empirical results, in that it is at this moment difficult to establish a comprehensive macro theory because such a theory would require first of all the comparing of micro models to find their underlying collective dynamics. Then a next step could be to find the additional ingredient of conditional information by excluding it from random noise and error.

3. Conditional Information in Social and Behavioral Research

In the sub-sections in this section, the focus is on the issue of situation in the actual world and the possibility of conditional information to reveal the connections between situation and person(s) as a means to obtain an understanding of when and why things happen in the actual world as they do.

3a. Literature Discussion

The developments in probability theory, system analysis theory, and social dynamics theory have provided for a better understanding of the challenges that exist regarding the detection of conditional information to clarify the connection(s) between situation and person(s). To summarize, probability theory holds a deterministic starting point, in that it is based on the assumption that events occur under certain conditions, in that knowledge of prior information can improve the calculations of probability. System analysis theory is aimed at explicating the functionality of an implementation strategy by analyzing empirical new data and employing simulation to validate the appropriateness of implementation strategies in various contexts. Social dynamics theory is based on the assumption that groups exhibit unique and emergent properties that cannot be understood comprehensively by merely aggregating the behavior of individuals. Regarding the detection of conditional information, these three theories all emphasize the need of assumption testing and they acknowledge that the theories can lead to general laws (i.e., causality) and feedback information that require further validation testing by including timescales, environment, and reactivity, while avoiding noise and error.

The literature overview also shows that (a) if not all prior information that is essential in a situation is included comprehensively (i.e., has become observable and knowledgeable) in data analysis, and it is acknowledged that (b) the dynamics of the situation hinders the detection and extraction of essential information, and (c) the presence of humans can further complicate situation interpretation by producing varied and nuanced behavior or actions, then the main difficulty for researchers is the validating of theories against empirical data. More specifically, research on detecting conditional information from situations in the actual world is largely dependent on observation possibilities and human interpretation regarding the analysis of empirical data to further theorizing.

Regarding human observation and interpretation in general, the human brain deals with actual-world information by simultaneously employing multiple mental processes that not only require time, in that the processes take place at the expense of continuous observation, but it also can be limited by the availability or accessibility of the memorized experiences and knowledge, and the appropriate processing of present emotions and thoughts (i.e., logical and critical reasoning and creative imagination). From an evolutionary perspective, humans attend to actual-world information if that can advance their survival and identity (see Chan, 2020, for a review on self-identity; see Slavich, 2020 for a biologically-health based review; see Windzio, 2023, for a social-psychologically based review). Others (Cohen & Marron, 2020) have argued that all universal adaptations of life are in the end driven by the universal properties of matter, which are energy (i.e., all forces that can create metabolic structures), entropy (i.e., all forces that can dismantle ordered structures), and interaction (i.e., inorganic and organic dynamics in the environment).

Regarding data observation and interpretation in research studies, to obtain a better understanding of why it is difficult to detect conditional information through empirical studies can be illustrated via the aforementioned examples of Brownian motion and food allergies. Brownian motion regarding colloids in fluids and gasses has been studied in two ways, namely as physics-theoretical formulations and empirical studies (see Bian et al., 2016; Haw, 2002; Philipse, 2018, for historic overviews). The physics-theoretical formulations, such as Einstein-Stokes general diffusion equation and Flick's diffusion law that reckon with environmental flux density and the colloids' characteristics, showed that diffusion further depends on the aforementioned (a) timescale (e.g., different timescales exist for particles and molecules), (b) environment (e.g., density, flow, and friction), and (c) reactivity (e.g., catalytic and transport colloidal interactions). The physics-theoretical formulations came forth from the first empirical studies of Brown (1828) that showed the existence of colloids, which were followed initially by conflicting and confusing reports regarding the external influences. Eventually, Perrin (1916) showed empirically that the random movement of spherical colloids in fluids and gasses depends on their kinetic energy (i.e., size and velocity) in relation to the energy dissipation of the surroundings, which determine the distance and the rate of diffusion of the colloids as a collective.

However, the consequences of different kinds of colloids in actual-world fluids and gasses are not yet understood completely because of observational issues (see Bian et al., 2016, for a discussion). That is, Brownian motion is not yet tractable in actual fluid and gas because observation requires a microscopic environment, and this influences how all the particles are in a particular fluid and gas, and how they interact among themselves and in relation to the present objects and barriers. For example, the currently often-employed method of three-dimensional digital video microscopy (3D-DVM) can provide for information on the diffusion of spherical polystyrene colloids in fluid. In either a plain microscope slide or a drilled one (i.e., having a cavity that allows for an air and water interface), and by adding ultra-purified water and a certain temperature, deionizing, NaCl levels, kinds of colloids, and low to high concentration of hydrodynamic interaction, the environment is created. Then the 3D-DVM can collect the ring changing profile of a colloid by doing several scans upward, and the computer calculates the objective movement of the colloid's position by using its diffraction pattern or brightness intensity. Similarly, the collection of multiple colloids enables the calculation of the mean movement of the colloids from the focal plane, resulting in a mean diffusion graph that visualizes the average distribution behavior of the colloids. Although 3D-DVM is successful in tracking individual and collective colloid motion, its limitations are that the colloids have to be well-defined and in good imaging conditions because the accuracy of the motion degrades in more complex environments (Fischer et al., 2011; Helgadottir et al., 2020).

Another example involves the behavior of non-spherical colloids in fluids and gasses. For instance, Doan et al. (2020) studied the Brownian motion of 3D polymeric colloids (i.e., 5 and 10 μm) that have spherical shapes in comparison to those with non-spherical shapes, namely polyhedrons (e.g., the square pyramid and tetrahedron), poly-pods (e.g., the tetrapod and hexapod), and complex polyhedral (e.g., the screw). The motion of the colloids was observed via confocal laser scanning microscopy. The results showed that the hindrance factor of the non-spherical shapes was higher compared to the spherically shaped colloids, leading to a slower diffusion rate of the first kinds of colloids. Because all colloids had an equivalent volume or mass, the hindrance factor is expected to come forth from the different amounts of drag at the substrate due to colloid shape (i.e., it was hypothesized that the hindrance factor is due to a larger contact area between colloid shape and substrate). Therefore, the empirical observation of Brownian motion does not yet account for colloids' actual-world behavior because the microscope confines the colloids between two glass plates. Moreover, actual colloid behavior is difficult to observe because (a) the colloids' behavior includes cluttering, merging, splitting, and et cetera among themselves and other particles and objects in the fluid or gas, and (b) there are signal detection issues with regard to parallel colloid trajectories and the colloids' signal intensity level in relation to the mean background noise level.

Similarly, research (see Lessens et al., 2020; Malone, 2024, for overviews) on food allergies and adverse reactions also provides for general information (i.e., food is the cause of an allergy because eliminating the food from the diet stops the allergy symptoms), which is roughly comparable to the general diffusion equation for Brownian motion because allergies are assessed via general testing criteria that produce certain information (see Spergel, 2019, for an overview). For example, the skin prick tests can assess reliably a specific food allergy, but they are not predictive of the severity of the allergy symptoms, which is mostly assessed via blood testing and food challenging that can be limited in assessing person-specific health complaints. Also comparable to research on Brownian motion in terms of providing for insufficient information about the contributing factors for having food allergies and adverse reactions are the issues of (a) cross-reactivity and environmental contributions to food allergy are not well known, (b) persons who test negative on skin tests still can have food allergy symptoms (i.e., non-IgE mediated allergies), and (c) food adverse reactions are difficult to assess. Then again, some participants' reports (see Rustad et al., 2022) about experiencing benefits after having eliminated certain foods from their diet imply that certain dietary changes can be benefiting. The present limitations in research on food allergies and adverse reactions showed that a more thorough examination is needed regarding the connections between foods and food components, the functioning of the immune system and the digestive system (e.g., enzymes, intestine and skin barrier function, and microbiome), and individual experiences and environmental influences during dietary changes. Overall, research on food allergies and adverse reactions showed that health problems also depend on (a) timescale (e.g., different timescales for food intake, digestion, and symptoms), (b) environment (e.g., food availability, eating pattern, and life style), and (c) reactivity (e.g., food components and interactions and cross-reactivity).

To conclude, the research on both Brownian motion and food allergies and adverse reactions suggests that important features that can support in detecting conditional information in situations are timescales, environment, and reactivity. Van Velzen (2023b) redefined these features regarding the social and behavioral sciences as timescales, context, and personal and situational circumstantial influences.

3b. Actual Situation, Person, and Conditional Information

The process of detecting conditional information for the social and behavioral sciences that can be inferred from the literature overview is twofold (see Figure 1). First, the definition of conditional information as a stipulation in terms of "This will happen *if* . . ., because . . ." can be useful for empirical research if it has the form of completing a general law and its details. Hence, conditional information has the form of "If Situation X_i takes place, because certain conditional information of . . . is present and influential, then the events of Y_i will occur or be experienced by certain persons because for them it means that . . ." The definition of conditional information now includes the stipulation of something being influential for a sub-class (i.e., objects, phenomena, and processes), for instance, in that it has a specific function for a certain situation in relation to certain persons in the plural. Second, the literature overview also implies that measuring the influence of situation and person requires the mapping out of the givens of the actual situation as comprehensively and realistically as possible because this can point out the variation that occurred within the situation (i.e., the sub-situations that have small changes in the events). Figure 1 illustrates this twofold process of detecting conditional information by presenting the example of identifying inter-individual groups based on person-specific intra-individual data, which is a growing research subject in the social and behavioral sciences where it defies that intra-individual variability is accidentally (Van Geert, 2012).

Situation-$_{average}$ in Figure 1 refers to a situation as it is experienced by everyone and everything that is present in that situation. This implies that it involves an overall, agreed-upon view and interpretation of all persons that are present and all of the features of the objects, phenomena, and processes as observed and experienced by all persons present (i.e., their knowledge about the present givens and the perspectives that they share). Where people agree on situations, they primarily refer to socially agreed upon knowledge and behavior, which in turn acknowledges certain situational features (Hünefeldt & Schlitte, 2018; Mousaïd et al., 2013), but their disagreements can remain obscured and, hence, reveal only part of the actual situation (Asendorpf, 1992; Molenaar, 2004; Van Geert & Van Dijk, 2021). The actual situation combines the objective situation, as agreed upon by everyone present in that situation, and the individual subjective interpretation of that situation, in that it can be argued that the actual situation has given the individual the possibility to come up with the subjective interpretation if it is real for several of the persons in that situation (Van Velzen, 2023a). Therefore, a difficulty of Situation-$_{average}$ is that it ignores individual views and interpretations (i.e., in the plural of several persons) that are part of the actual situation. These individual views and interpretations are based on individual person's characteristics (i.e., knowledge about (a) familiar information about actual-world objects, phenomena, and processes, including oneself as a person, (b) information necessary to execute complex reasoning, and (c) information about actual-world changes as a possible consequenc of the passing of time, including expected social behavior: Van Velzen, 2022b).

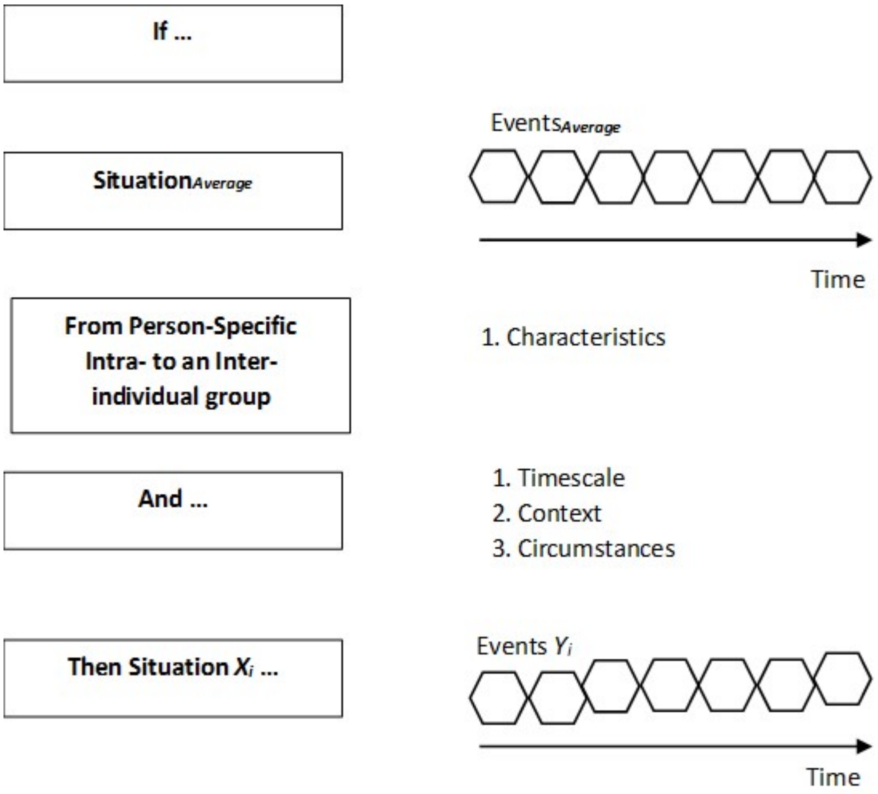

Figure 1. Schematic presentation of the features involved for detecting conditional information in actual-world situations.

Accordingly, when an individual person's situational subjective observation and interpretation are taking into account by means of their personal characteristics, then a similar situational description from several persons regarding the conditional information in relation to timescale, context, and circumstances (i.e., inferred from literature review) can point out an agreed upon sub-situation. In Figure 1, the agreed upon sub-situation that is realistic for a group of individuals is schematically presented by sub-situation X_i that has slightly varying events in comparison to the events of Situation-$_{average}$. Hence, the thus identified inter-individual group matches the description of "Situation X_i occurs with the events of Y_i." For example, if homework learning (i.e., studying and memorizing) takes place (i.e., Situation X), in that a sample of students describe their learning during homework, and several students in the sample describe their homework learning as for them it is essential to learn-for-understanding (i.e., it has an influence), then the learning events (i.e., Y_1) that this group of students describe is a certain kind of homework learning. Hence, these students that have this kind of homework learning become the inter-individual group of viewing and interpreting homework-learning situations as Situation X_1 (i.e., learning-for-understanding).

Specifically, this inter-individual group that views the homework-learning situation slightly different from the average agreed-upon views on the homework-learning situation is illustrated in Figure 1 by the events Y_i of the inter-individual group going slightly upwards in direction in comparison to the average agreed-upon events. That is, Figure 1 shows two kinds of events-in-a-situation, namely the straight horizontal line of events in Situation$_{Average}$, which indicates that the average situation of all collected, latent, and emerging data is taken as the point of departure for what the situation encompasses, and the slightly upward trend of the events in Situation X_i denotes its deviation from Situation$_{Average}$. Notably, the slightly upward trend of events illustrated in Figure 1 as Situation X_i is only an example, in that other directions of deviation regarding the straight horizontal line of Situation$_{Average}$ are also imaginable, such as a slightly downward trend of the events. Having a changing pattern of events in the various sub-situations of X_i is helpful not only to establish inter-individual groups from person-specific intra-individual data, but, as Van Velzen (2022a, 2023a) argued, it is also better in line with the dynamics of situations in the actual world, the relativity of the observers of situations, and the passing of time. That is, the dynamics of situation as they appear in the actual world refers to the ever-continuing passing of time via changes in actual-world states, in that the objects, phenomena, and processes that make up the actual-world states produce patterns of change when states make up events that continuously go from a near-past to a present time. Hence, the continually changing actual world can show degrees of changes regarding the patterns of behavior when the continuing affair of a situation is unraveled in several sub-situations and their events and how this can lead to changes in individual views and interpretations. This means that reckoning with timescales as temporal window frames (Buehner, 2005) can be helpful for identifying sub-situations.

Person-specific intra-individual data are thought to be essential for identifying inter-individual groups rather than believing that the intra-individual variability that is found in empirical studies is accidentally (Van Geert, 2012). Although identifying inter-individual groups based on all-including (i.e., original without data cleaning) person-specific intra-individual data seem sensible, it is also complex, from a data-analytical point of view, to obtain an understanding of how individual behavioral trajectories can contribute to identifying inter-individual groups (Asendorpf, 1992; Molenaar, 2004). First, it requires the enabling of collecting emerging data because this kind of data can contribute to obtaining an understanding of individual behavioral trajectories in terms of providing for the missing link(s) regarding the intentionally collected data (Glaser & Strauss, 1967). Specifically, the collecting of emerging data are likely obtainable best via qualitative data-collection instruments even when this may complicate validation (Bendassolli, 2013). To this end, Van Velzen (2018; 2020) argued that because useful emerging data can consist of personal experiences, mixed methods research data collection could produce integrated qualitative data. That is, obtaining emerging data via mixed methods research can eliminate the uncertainties that may arise from heterogenic data because integrated data can embrace complexity in a more holistic manner (Budischak et al., 2018; Mast et al., 2014; Sears et al., 2011; see also Van Velzen, 2021, for an overview).

Mixed methods research data collection can involve quantitative or numerical data, quantitized or qualitative data that enable a hierarchical and ordered numerical translation (i.e., nominal, ordinal, and interval measurement scales), qualitative or verbal data that are collected via open-ended instruments and observed behavior, and if possible, the construction of new sets of consolidated data (see Van Velzen, 2018), because each can provide for information about the objective components in actual-world situations (e.g., location, context as objective overall act, situation, events, features, and moments) and the subjective components (e.g., perspective, focus, and position: see Van Velzen, 2023a). For example, the objective situation theoretically includes (a) the amount, kinds, and spatial position of objects as they are present in the situation, (b) the overall act of what is happening in terms of everyone agreeing about a series of subsets and events, (c) their physical elements, such as the amount and the surroundings of persons, temperature, and conversations, and (d) the changes that take place at certain moments in time (see Figure 2).

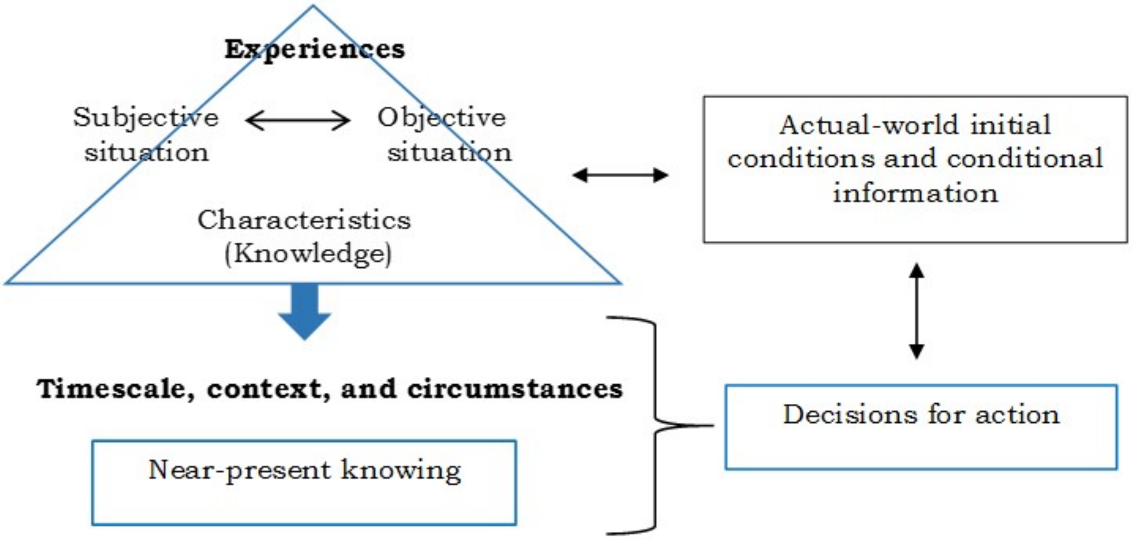

Figure 2. Theoretical framework providing a schematic representation of actual world situations as their included information across time (in black) in relation to human interpretation (in blue). See Van Velzen, 2024b.

This is also called the givens, in that the givens represent all that is present in the situation. The subjective experiences that are memorized by a person involves this person’s knowledge about oneself and the world, and the subjective situation involves the person’s perspective (i.e., physical-spatial position), direction and intensity of attention, and interpretation of timescales, context, and circumstances (i.e., including thoughts and feelings). Especially the emerging data are thought to be of interest to collect information about the difference between objective and subjective experienced situation because the emerging data can produce better-integrated meanings of what is actually happening in a situation compared to the overall mean of agreed-upon meanings that are often lacking in explanations.

There is also one important limitation, namely how to see the forest for the trees in the data when they consist of quantitative, quantitized, and qualitative data as well as latent and emerging variables and information. The main argument against the collecting of mixed methods research person-specific intra-individual data is that the amount of variables are likely to produce disorganization and contradictory findings and results. Then again, it also matters how the data are analyzed and, of course, data analysis will require an open-minded attitude of seeing whether it will add to rather than distort what one was expecting to find, and this may not always be easy to do because it can be challenging to conduct mixed methods research studies (see Adu et al., 2022, for an overview).

Notably, the argument against the collection of mixed methods research person-specific intra-individual data because of the amount of variables and the resulting disorganization and contradictory findings and results has nothing to do with the limitations often ascribed to big data analysis because the data employed for big data analysis often come from multiple survey cohorts and, therefore, are not collected for one specific research question, and big data can be unsuitable in quality, scalability, and integration (Hofferth et al., 2017; Tosi et al., 2024). For example, survey cohorts often inquire about the so-called “easy” to collect, quantitative information, such as weights, years, temperature, centimeters, and et cetera, that even when these kinds of data are taken together, it often cannot provide for a holistic interpretation (c.f., Sears et al., 2011).

Overall, the amount of variables in mixed methods research person-specific intra-individual data analysis can become extensive and overwhelming, especially, when it is also accounted for that the data will include the complexity and dynamics of situations across time. Therefore, it is legitimate to raise the question of how to deal with this kind of data via a data-analytical technique, which will be the subject of the next section.

4. Initial, All-Including, and Visualized Data Analysis

The sub-sections in this section propose a specific data-analytical technique for the purpose of identifying inter-individual groups based on person-specific intra-individual data via the detection of conditional information because it is suitable for initial exploration and description of all-including mixed methods research data (for a preliminary illustrative example of this technique, see Van Velzen, 2024b).

4a. Four-Quadrant Relationship Data Analysis

The main difficulty in mixed methods research data analysis of person-specific intra-individual data is the coming into existence of emerging data that are produced, in particular, by the qualitative data. The emerging of data in a research study can enlighten and confuse with regard to making informed choices about the meaning of the data because it may and may not increase the complexity (i.e., via an abundance of information) and dynamics (i.e., via varying patterns of change). The collection of mixed methods research data might lead to obtaining more and different kinds of extreme scores, which can cloud the theory. More and different kinds of extreme scores can complicate data analysis because assuming a Gauss distribution regarding the obtained scores might not be self-evident and the data might be interdependent. Conversely, obtaining information or insights via qualitative and emerging data about *why* participants make certain (behavioral) decisions regarding specific situations can clarify the theory, for instance by providing for conditional information, because this information can explain certain situational influences on (behavioral) decisions. Therefore, it would be prudent if data analysis could enable a thorough understanding of all data and the detection of conditional information in terms of connecting all the quantitative, quantitized, and qualitative person-specific intra-individual data, as well as latent and emerging data.

Regarding the social and behavioral sciences, data analysis should enable a thorough overview per individual participant not only to obtain an understanding of individual decisions, behavior, and actions, but also to identify how several individual persons can be related to one another regarding their views on and interpretations of a situation. When persons, who are in agreement about the kinds and values of certain information in a situation can form an inter-individual group, then this information can specify certain influential information as conditional information.

As a practical and initially essential working method of data analysis for the purpose of the complex problem solving that is the identification of inter-individual groups based on person-specific intra-individual data regarding a situation, obtaining a comprehensive and descriptive overview of all data is essential because in this way the researcher can make sure that all available information in the data is considered with regard to whether it is meaningful or functional for certain groups of people in a situation. Commonly, data are checked first for the purpose of exploring the structure and identifying trends and clusters, for instance, by applying the standard statistical techniques for obtaining the mean (e.g., median, mode, and skewedness), variance and standard deviation, and correlation and regression coefficients for the quantitative and quantitized data. To support in interpreting the thus obtained general data overview, visualization can be established by plotting histograms, diagrams, and matrixes.

A specific kind of data visualization, the so-called the four-quadrant plot is employed commonly to show simultaneously the relationship between two kinds of scores for negative and positive values. For example, the four-quadrant plot (see Figure 3a) can be applied to display the relationship between two methods (e.g., the test method and a reference method) at various time plots via the direction of change of delta-scores of both methods to determine the concordance (c.f., Critchley et al., 2010). This kind of application of the four-quadrant plot is used especially in cardiac output monitoring research (Bataille et al., 2012; Proença et al., 2019), but also regarding individual cognitive variability, for instance, in the perceptual load during information processing and decision making (Marciano & Yeshurun). In addition, the four-quadrant plot is also employed for qualitative data to distinguish thematic distinctions (Jang et al., 2021), and then it does not necessarily have to ensure a division for positive and negative values, it only has to distinguish the thematic distinctions.

However, all standard statistics, including the four-quadrant plot for positive and negative values, assume a stationary mean (i.e., a constant mean) and not a time varying mean across situations (c.f., Molenaar & Lawrence, 2012). To overcome this issue, a visualization of numeric and verbal or thematic data simultaneously can be achieved by the so-called four-quadrant relationship display (Van Velzen, 2024b), in that it can position individual participant's quantitative, quantitized, and qualitative values and concepts, as well as latent and emerging data, and across time (see Figure 3b). By placing each individual's score in the four-quadrant relationship display to visualize their position regarding two (or more) variables, values, and themes, those individuals with similar positions in one of the four quadrants can be detected and, the analysis of multiple four-quadrant relationship displays can then identify inter-individual groups. In this way, the identification of inter-individual groups is based on all data. Furthermore, the multiple four-quadrant relationship displays can ascertain an understanding of why a particular inter-individual group was identified as a group of individuals when the identification of the inter-individual group involves conditional information.

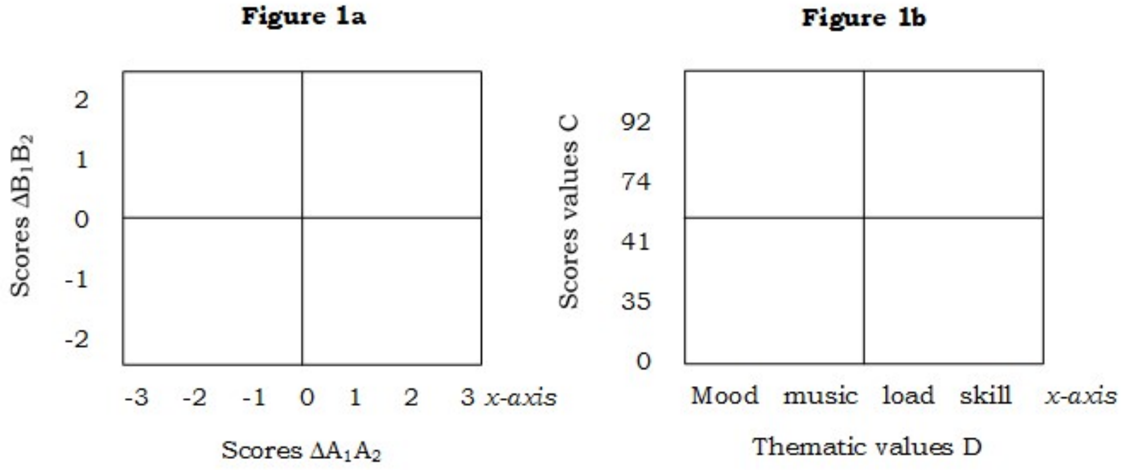

Figure 3. Schematic presentation of the mathematical four-quadrant plot (a) and the proposed four-quadrant relationship display (Van Velzen, 2024b).

Figure 3b shows that qualitative data, such as certain words, related groups of words, and words that are expressing certain emotional behavior, which are placed on the *x*-axis, can be compared to the quantitative and quantitized data, that are placed on the *y*-axis either separately or as consolidated new data sets. Of course, the construction and comparing of multiple four-quadrant relationship displays is intensive in terms of the amount of workload required, but it is a sensible manner for initiating data analysis to obtaining a thorough overview of all data and it can deal with complex information about situation in relation to individual persons and across time.

To clarify why a four-quadrant relationship display is sensible for initial mixed methods research data analysis and for obtaining a thorough understanding of all-including data, Unwin's (2020) overview of data visualization presents arguments regarding the importance of visualizations especially for exploratory data analysis, namely that visualizing can display data positions (i.e., of all-including data), enlighten complicated transformations, and reveal data structures and features that statistics and models may miss. Exploratory graphical displays are useful to find and direct attention to the information in the data, in that they can support thinking. Unwin (2020) also discusses research on new and innovative graphical displays that, in the future, may facilitate the interpretation of multiple graphical displays by combining the information of several graphical displays into one, such as the mosaic plot. Therefore, appropriate visualizations of data can reveal more information about all-including mixed methods research data then either standard statistics or words and themes can.

However, does this mean that there are no drawbacks and limitations regarding exploratory visualized data analysis? In general, visualizations can provide for accuracy in decision making, if the (a) quality is high of both the data and the created graphical display to present the information comprehensively and (b) interpretation of the graphical display is made critically because graphical displays are being devised to improve critical thinking (Eberhard, 2023; Chen et al., 2009). This raises the question of when exploratory visualized data analysis, in general, and the four-quadrant relationship display, specifically, can lead to contradictory interpretations. Overall, contradictory interpretations are to be expected when there are insufficient data available and when the data are inappropriate. For example, the identification of inter-individual groups requires information about the persons in the situation, such as the way in which individual persons and a group(s) of persons react in the situation regarding one another, use (non-) verbal and emotional behavior, and deal with roles and settings. Insufficient and inappropriate data are a limitation that is not restricted to visualization of exploratory data analysis because it is always essential to collect sufficient and appropriate data, but a large amount of varying data can hinder standard data analysis more easily.

Another question is whether the identified inter-individual groups via four-quadrant relationship data analysis are valid. Regarding the validity of the four-quadrant *plot*, this is found to be well-grounded for the comparing of cardiac output monitoring tests (Bataille 2012), but to my knowledge, there are no results regarding the validity of the four-quadrant relationship data-analysis and its application to identify inter-individual groups based on all-including, person-specific intra-individual data. Validation or obtaining knowledge about the accuracy of the interpretations made based on data (Messick, 1995) is important to increase the trustworthiness of the results obtained via four-quadrant relationship displays. In this respect, especially, the contextual differences found in various situations can inform about the accuracy of the identified inter-individual groups (Adcock & Collier, 2004). Van Velzen (2020; 2023b) argued that contextual differences across various situations involve the global and local environment in relation to the objective and subjective situational values and features, as they go from one near-situational moment (i.e., having various timescales and personal and situational circumstances) to the next near-situational moment.

There is preliminary information regarding the feasibility of the four-quadrant relationship display as a data-analytical technique, in that the illustrative example (Van Velzen, 2024b) showed that the four-quadrant relationship display is suitable for obtaining an initial data overview of all-including data, an understanding of the data, and it can connect quantitative, quantitized, and thematic qualitative data. In the illustrative example, the four-quadrant relationship data analysis enabled the identification of inter-individual groups based on all-including person-specific intra-individual data and it led to the detection of conditional information. Unfortunately, the validity of the identified inter-individual groups and the detected conditional information were not established because a limitation of the illustrative example was that the data were not collected specifically for studying the detection of conditional information.

4b. Four-Quadrant Relationship Data Analysis for Conditional Information

The importance of enabling data to emerge via data collection and to analyze the emerging data in order to identify inter-individual groups based on person-specific intra-individual data has been emphasized throughout this article. Regarding the detection of conditional information as a means to connect situation and persons, as in groups of comparable persons in the way they view and interpret a situation, the importance of analyzing emerging data where it concerns exploratory data analyses has been emphasized because when nothing is known about conditional information in certain situations for groups of persons, it cannot be conceptualized in variables and operationalized via measurements. It has also been argued that emerging data are obtained best via qualitative data collection. However, analyzing qualitative data can be a daunting task because of the complexity of inferring the precise and correct person-specific meanings from open-ended responses, to which end the data collection needs to elicit meaningfully rich and comprehensive information (c.f., Pawson & Tilley, 2014). This raises the question of what are the latest techniques in qualitative and exploratory data analysis and visualization.

Standard qualitative data analysis consists of content analysis to infer themes and categories, in that it requires meaningful and critical inquiry, contextualization, and interpretation. Thorough explorative qualitative data analysis can stimulate our thinking in order to understand a research subject in some meaningful manner, for instance, about the similarities and differences among people's experiences in and of situations (Thorne, 2020) and, specifically, by including emerging data in the data analysis to evaluate all data even though they are different or do not fit (Eakin & Gladstone, 2020). Currently, computational text analysis can explore written texts in natural language to extract structures and features, which is advantageous especially when the data consist of large and unstructured textual information (Figura et al., 2023; Ke et al., 2024). Another advantage of computational text analysis is that the coding of text can be more objective than standard content analysis, and it allows for statistical calculations. However, a limitation of computational text analysis is that it is restricted to identifying sameness or similar words rather than evaluating that which is different, in that it can support researchers in content analysis if they add further critical analysis. Overall, computational text analysis always needs further standard content analysis, to which end visualizations can be helpful by organizing, reducing, and discussing the qualitative data (Watkins, 2017). Of course, inferences made on the basis of the obtained data need to be established carefully, in that putting the inference process down into writing is difficult, but essential in order to understand how the process has led to an inference and a certain meaning (see Bischoping, 2005, for a discussion).

Regarding the analyzed qualitative data, as aforementioned, the advantage of the four-quadrant relationship data analysis is that it can connect qualitative data to quantitative and quantitized data. In the illustrative example (Van Velzen, 2024b), the qualitative data produced the quantitized data and it was essential to enable the *final* data analysis that connected the summary of several four-quadrant relationship displays to understand why the identified inter-individual group was a group. Something similar was pointed out by Bischoping (2005), who argued that the solitary part of going through the qualitative data, over and over, is essential to analyze the data deeply and thoroughly to establish justifiable interpretations, in that the end result shows an increase in the accuracy of the interpretations made. Analyzing all-including data to understand it thoroughly by going over and over it, can be said also about evaluating multiple four-quadrant relationship displays because of its importance for making justifiable interpretations.

When all data are placed into multiple four-quadrant relationship displays, then a next question is if this can detect conditional information. The illustrative example (Van Velzen, 2024b) showed the presence of both initial conditions and conditional information. Initial conditions, defined as representing a constant or starting value, were found in the illustrative example on homework learning as follows. Learning differs per subject, in that, for instance, learning for a language differs from learning for mathematics. Because most participants (78%) mentioned this feature regarding their homework learning in the open-ended responses, it can be included in further data analysis as an initial condition. Regarding the conditional information found in the illustrative example, those participants that mentioned to learn-for-understanding during homework learning formed an inter-individual group. The conditional information, defined as additionally relevant information that provides for a stipulation that is essential to a group of students, was found and can be considered as influential because the individual students mentioned it of their own accord and independent of one another.

This raises the question of whether conditional information can be detected via timescale, context and circumstances as assumed in the theoretical framework (see Figure 2). First, the results of the illustrative example imply that the objective situation refers to the situation as it is in agreement for all participants and it can produce information about initial conditions, whereas the subjective situation is the situation as it agrees with an identified group of participants, and this can produce information about conditional information for this group of participants. Second, the results of the illustrative example provided for very general information about timescales, context, and circumstances in relation to the detection of conditional information. Then again, the data for the illustrative example was not collected specifically to detect conditional information. Examples involving context are “If I have to study very complex subject matter, then I will study until I understand it all, unless I become tired and stressed out for exams” (i.e., without going into the kinds of contexts that produce tiredness and exam stress), and “I never study longer to improve my learning, but I do take a look at and analyze previously made tests” (i.e., without explicating the context by mentioning when and why previous tests are reexamined). Hence, more research is needed to answer this question about detecting conditional information regarding situation and persons, as well as related questions, such as whether the study of the objective and subjective situational features should require multiple moments of data collection (e.g., time series) to understand better the process of situational change and the function of certain conditional information. Based on this theoretical article and the illustrative example, conditional information for the social and behavioral sciences can be defined now and preliminarily as follows.

Conditional information defined as additionally relevant information, in that it provides for a provision or stipulation, refers to information that is found via all-including data and that is meaningful and influential (i.e., describing the when and the why) in a sub-situation (e.g., of an inter-individual group) of the research subject under investigation.

Conclusion

When there is a situation that consists of a series of events and states, and this situation includes objects, phenomena, and processes that together produce an affair that evolves across time, then the result will consist of changing patterns in behavior, events, and trajectories. In empirical studies, some of these objects, phenomena, and processes are observed and others are either latent and emerging or not becoming available for unknown reasons. Then there are also error and noise that can further cloud that which is observed in relation to that which is actually happening. From a data-analytical perspective, causal inference, directionality in influence, and conditional expectations obtained via (non-) linear regression and other multivariate data-analytical techniques may not describe and explain easily what is actually happening in a situation because it describes and explains only part of the actual situation by assuming a constant mean. This argues in favor of employing all-including initial data analysis to obtain a thorough overview of all the information that resides in the data. In this article, the four-quadrant relationship display is proposed as an initial data-analytical technique to enable such an overview because it can visualize not only the observed, latent, and emerging data, but also the connections between quantitative, quantitized, and qualitative data. Moreover, the detection of conditional information can provide for the missing link between generalized and detailed research results because it can produce otherwise overlooked connections, such as the identification of inter-individual groups based on person-specific intra-individual data. Finally, besides research on obtaining more information about the possibilities and challenges of the four-quadrant relationship data analysis, for it to be a customer friendly data-analytical technique, more advanced computer programs are required that can connect multiple four-quadrant displays easily into overviews.

References

Achterhold, J., & Stückler, J. (2021, April). *Explore the context: Optimal data collection for context-conditional dynamics models*. Paper presented at the 24th International Conference on Artificial Intelligence and Statistics, San Diego, California. Https://explorethecontext.is.tue.mpg.de.

Adcock, R., & Collier, D. (2004). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review, 95*(3), 529-546.

Adu, J., Owusu, M. F., Martin-Yeboah, E., Gavidia, L. A. P., & Gyamfi, S. (2022). A discussion of some controversies in mixed methods research for emerging researchers. *Methodological Innovations, 15*(3), 321-330.

Alon, U. (2007). *An introduction to system biology: Design principles of biological circuits*. Chapman & Hall/CRC.

Asendorpf, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality, 6*(2), 103-117.

Bartoš, F., Sarafoglou, A., Godmann, H. R., Sahrani, A., Leunk, D. K., Gui, P. Y., ..., & Wagenmakers, E. J. (2023). Fair coins tend to land on the same side as they started: Evidence from 350,757 flips. *ArXiv Mathematics, History and Overview*: 2310.04153v2.

Bataille, B., Bertuit, M., Mora, M., Mazerolles, M., Cocquet, P., Masson, B., ..., & Larché, J. (2012). Comparison of esCCo and transthoracic echocardiography for non-invasive measurement of cardiac output intensive care. *British Journal of Anaesthesia, 109*(6), 879-886.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances: By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, M. A. and F. R. S. *Philosophical Transactions of the Royal Society of London, 53*(1), 370-418.

Bechinger, C., Di Leonardo, R., Löwen, H., Reichhardt, C., Volpe, G., & Volpe, G. (2016). Active particles in complex and crowded environments. *Review of Modern Physics, 88*(4): 045006.

Bendassolli, P. F. (2013). Theory building in qualitative research: Reconsidering the problem of induction. *Forum: Qualitative Social Research, 14*(1). Https://doi.org/10.17169/FQS-14.1.1851.

Bian, X., Kim, C., & Karniadakis, G. E. (2016). 111 years of Brownian motion. *Soft Matter, 12*(30), 6331-6346.

Bischoping, K. (2005). Quote, unquote: From transcript to text in ethnographic research. In D. Pawluck, W. Schaffir, & C. Miall (Eds.), *Doing ethnography* (pp. 141-154). Canadian Scholar's Press.

Blesch, K., Watson, D. S., & Wright, M. N. (2023). Conditional feature importance for mixed data. *Advances in Statistical Analysis, 107*(3): 00477-9.

Borsboom, D., Deserno, M. K., Rhemtulla, M., Epskamp, S., Fried, E. I., McNally, R. J., ..., & Waldorp, L. J. (2021). Network analysis of multivariate data in psychological science. *Nature Reviews Methods Primers, 1*(1): 58.

Brown, R. (1828). A brief account of microscopical observations made in the months of June, July and August 1827 on the particles contained in the pollen of plants, and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine Series 2, 4*(21), 161-173.

Budischak, S. A., Hansen, C. B., Caudron, Q., Garnier, R., Kartizinel, T. R., Pelczer, I., ..., & Graham, A. L. (2018). Feeding immunity: Physiological and behavioral responses to infection and resource limitation. *Frontiers in Immunology, 8*: 01914.

Buehner, M. J. (2005). Contiguity and covariation in human causal inference. *Learning & Behavior, 33*(2), 230-238.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *ArXiv Physics, Social Philosophy*: 0710.3256v2.

Chan, K. T. (2022). Emergence of the 'digitalized self' in the age of digitalization. *Computers in Human Behavior Reports, 6*(1): 100191.

Chen, M., Ebert, D., Hagen, H., Laramee, R. S., Van Liere, R., Ma, K. L., …, & Silver, D. (2009). Data, information, and knowledge in visualization. *IEEE Computer Graphics and Applications, 29*(1), 12-19.

Choi, M., & Lee, C. (2024, May). *Conditional information bottleneck approach for time series imputation*. Paper presented at the Twelfth International Conference on Learning Representations, Vienna, Austria. Https://openreview.net/pdf?id=K1mcPiDdOj.

Cohen, I. R., & Marron, A. (2020). The evolution of universal adaptations of life is driven by universal properties of matter: Energy, entropy, and interaction. *F1000 Research, 9*: 626.

Covert, I., Lundberg, S. M., & Lee, S. I. (2020). Understanding global feature contributions with additive importance measures. *Advanced in Neural Information Processing Systems, 33*(1), 17212-17223.

Critchley, L. A., Lee, A., & Ho, A. M. H. (2010). A critical review of the ability of continuous cardiac output monitors to measure trends in cardiac output. *Anesthesia & Analgesia, 111*(5), 1180-1192.

De Finetti, B. (2017). *Theory of probability: A critical introductory treatment*. Wiley.

Doan, D., Kulikowski, J., & Gu, X. W. (2020). Diffusion of anisotropic colloidal microparticles fabricated using two-photon lithography. *Particle & Particle Systems Characterization, 38*(8): 21000333.

Docquier, D., Di Capua, G., Donner, R. V., Pires, C. A. L., Simon, A., & Vannitsem, S. (2024). A comparison of two causal methods in the context of climate analyses. *Nonlinear Processes in Geophysics, 31*(1): 115-136.

Eakin, J. M., & Gladstone, B. (2020). “Value-adding” analysis: Doing more with qualitative data. *International Journal of Qualitative Methods, 19*(1), 1-13.

Eberhard, K. (2023). The effects of visualization on judgment and decision-making: A systematic literature review. *Management Review Quarterly, 73*(1), 167-214.

Figura, M., Fraire, M., Durante, A., Cuoco, A., Arcadi, P., Alvaro, R., …, & Piervisani, L. (2023). New frontiers for qualitative textual data analysis: A multimethod statistical approach. *European Journal of Cardiovascular Nursing, 22*(5), 547-551.

Fischer, R. S., Wu, Y., Kanchanawong, P., Shroff, H., & Waterman, C. M. (2011). Microscopy in 3D: A biologist’s toolbox. *Trends in Cell Biology, 21*(12), 682-691.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine Transaction.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424-438.

Granger, C. W. J., & Watson, M. W. (1984). Time series and spectral methods in econometrics. In Z. Griliches & M. D. Intriligator (Eds.), *Handbook of econometrics* (Vol. 2, pp. 980-1022). Elsevier.

Harton, H. C., & Bullock, M. (2007). Dynamic social impact: A theory of the origins and evolution of culture. *Social and Personality Psychology Compass, 1*(1), 521-540.

Haw, M. D. (2002). Colloid suspensions, Brownian motion, molecular reality: A short history. *Journal of Physics: Condensed Matter, 14*(33), 7769-7779.

Helgadottir, S., Argun A., & Volpe, G. (2019). Digital video microscopy enhanced by deep learning. *Optica, 6*(4), 506-513.

Hofferth, S. L., Moran, E. F., Entwisle, B., Aber, J. L., Brady, H. E., Conley, D., …, & Hubacek, K. (2017). Introduction History and motivation. *The Annals of the American Academy of Political and Social Science, 669*(1), 6-17.

Hünefeldt, T., & Schlitte, A. (2018). Introduction: Situatedness and place. In T. Hünefeldt & A. Schlitte (Eds.), *Situatedness and place: Multidisciplinary perspectives on the spatio-temporal contingency of human life* (pp. 1-18). Springer.

Jang, F. L., Chien, T. W., & Chou, W. (2023). Thematic maps with scatter and 4-quadrant plots in R to identity dominant entities on schizophrenia in psychiatry since 2017: Bibliometric analysis. *Medicine, 102*(46): e36041.

Jaynes, E. T. (2018). *Probability theory: The logic of science*. Cambridge University Press.

Ke, Z. T., Ji, P., Jin, J., & Li, W. (2024). Recent advances in text analysis. *Annual Review of Statistics and Its Application, 11*(1), 347-372.

Kim, B., Cruden, G., Crable, E. L., Quanbeck, A., Mittman, B. S., & Wagner, A. D. (2023). A structured approach to applying system analysis methods for examining implementation mechanisms. *Implementation Science Communications, 4*(1): 127.

Knuth, K. H. (2016). *A modern history of probability theory*. Max Planck Institut für Extraterrestrische Physik. Https://www.mpe.mpg.de/~aws/knuth-mpi-talk---final.pdf.

Kolmogorov, A. N. (1933: English translation 1950). *Foundations of the theory of probability*. Chelsea Publishing.

Lei, J., G'Sell, M., Rinaldo, T., Tibshirani, R. J., & Wasserman, L. (2018). Distribution-free predictive inference for regression. *Journal of the American Statistical Association, 113*(523), 1094-1111.

Lessens, D., Gundlach, K., & Pollack, S. (2020). *Elimination diets*. University of Wisconsin-Madison: School of Medicine and Public Health. Https://www. famed.wisc.edu/integrative.

Libchaber, A. (2019). From biology to physics and back: The problem of Brownian movement. *Annual Review of Condensed Matter Physics, 10*(1), 275-293.

Malone, J. C., & Daley, S. F. (2024). *Elimination diets*. StatPearls Publishing.

Marciano, H., & Yeshurun, Y. (2017). Large inter-individual and intra-individual variability in the effect of perceptual load. *PLoS ONE, 12*(4): e0175060.

Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and Social Psychology Review, 11*(3), 279-300.

Mast, F. D., Ratushny, A. V., & Aitchison, J. D. (2014). Systems cell biology. *Journal of Cell Biology, 206*(6), 659-706.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201-218.

Molenaar, P. C. M., & Lawrence, L. L. (2012). Dynamic factor analysis and control of developmental processes. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 333-349). The Guilford Press.

Molnar, C., König, G., Bischl, B., Casalicchio, G. (2023). Model-agnostic feature importance and effects with dependent features: A conditional subgroup approach. In M. Atzmüller, J. Fürnkranz, T. Kliegr, & U. Schmid (Eds.), *Data mining and knowledge discovery* (pp. 1-39). Springer.

Mousaïd, M., Kämmer, J. E., Analytis, P. P., & Neth, H. (2013). Social influence and the collective dynamics of opinion formation. *PLoS ONE, 8*: e78433.

Pawson, R., & Tilley, N. (2014). *Realistic evaluation.* Sage.

Peralta, A. F., Kertész, J., & Iñiguez, G. (2022). Opinion dynamics in social networks: From models to data. *ArXiv Physics, Social Philosophy*: 01322v4.

Perrin, J. (1916). *Atoms*. D. van Nostrand Company.

Philipse, A. P. (2018). *Brownian motion: Elements of colloid dynamics*. Springer.

Proença, M., Bonnier, G., Ferrario, D., Verjus, C., & Lemay, M. (2019, July). *PPG-based blood pressure monitoring by pulse wave analysis: Calibration parameters are stable for three months.* Paper presented at the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society in Berlin, Germany.

Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos, 28*(7): 075310.

Rustad, A. M., Nickles, M. A., Bilimoria, S. N., & Lio, P. A. (2022). The role of diet modification in atopic dermatitis: Navigating the complexity. *American Journal of Clinical Dermatology, 23*(1), 27-36.

Sears, M. W., Raskin, E., & Angilletta, M. J. (2011). The world is not flat: Defining relevant thermal landscapes in the context of climate change. *Integrative and Comparative Biology, 51*(5), 666-675.

Sebo, S., Stoll, B., Scassellati, B., & Jung, M. F. (2020). Robots in groups and teams: A literature review. *Proceedings of the ACM on Human-Computer Interaction 4*(2): 176.

Slavich, G. M. (2020). Social safety theory: A biologically based evolutionary perspective on life stress, health, and behavior. *Annual Review of Clinical Psychology, 16*(1), 265-295.

Spergel, J. (2019). *Diagnosing food allergies in infants and children*. Pediatric Nutrition: Continuing Education for Clinicians.

Szabo, R. O., Chowdhary, S., Deritei, D., & Battiston, F. (2022). *Nature Scientific Reports, 12*: 10498.

Thorne, S. (2020). Beyond theming: Making qualitative studies matter. *Nursing Inquiry, 27*(1): e12343.

Tosi, D., Kokaj, R., & Roccetti, M. (2024). 15 years of big data: A systematic literature review. *Journal of Big Data, 11*: 73.

Unwin, A. (2020). Why is data visualization important? What is important in data visualization? *Harvard Data Science Review, 2*(1). Https://doi.org/10.1162/99608f92.8ae4d525.

Van Geert, P. (2012). Dynamic systems. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 725-741). The Guilford Press.

Van Geert, P., & Van Dijk, M. (2021). Thirty years of focus on individual variability and the dynamics of processes. *Theory & Psychology, 31*(3), 405-410.

Van Velzen, J. H. (2018). Students' general knowledge of the learning process: A mixed methods study illustrating integrated data collection and data consolidation. *Journal of Mixed Methods Research, 12*(2), 182-203.

Van Velzen, J. H. (2020). *Scientific reasoning in the face of struggling to further research: A critical search*. Sigmetack Publishing.

Van Velzen, J. H. (2021). Towards a working definition of experiences across time. *RRREaT-PT, Editorial 2021*(1), 1-9.

Van Velzen, J. H. (2022a). Theorizing about situations across time: The dynamics of the actual world. *RRREaT-PT, Editorial 2022*(1), 1-14.

Van Velzen, J. H. (2022b). What is knowledge as an actual-world phenomenon? *RRREaT – Cognitive Psychological Phenomena in Education, 1*: 01.

Van Velzen, J. H. (2023a). The multi-cube: The relativity of the observer of person in situation. *RRREaT-PT, Editorial 2023*(1), 1-23.

Van Velzen, J. H. (2023b). What is the actual world as the study object of scientific research? *RRREaT – Cognitive Psychological Phenomena in Education, 1*: 02.

Van Velzen, J. H. (2024a). Research and experiences across time: Exploring avenues. *RRREaT-PT, Editorial 2024*(1), 1-7.

Van Velzen, J. H. (2024b). Four-quadrant relationship data analysis: An illustrative example. *RRREaT – Data-Analytical Techniques, 1*: 02.

Watkins, D. C. (2017). Rapid and rigorous qualitative data analysis: The “RADaR” technique for applied research. *International Journal of Qualitative Methods, 16*(1), 1-9.

Warren, W. H., Falandays, J. B., Yoshida, K., Wirth, T. D., & Free, B. A. (2024). Human crowds as social networks: Collective dynamics of consensus and polarization. *Perspectives on Psychological Science, 19*(2), 522-537.

Windzio, M. (2023). The evolution of human sociality: Categorizations, emotions, and friendship. *Cologne Journal of Sociology and Social Psychology*. Https://doi.org/10.1007/s11577-023-00919-x.

Footnotes

[1] A discrete time series is here defined as a vector x_t of observations made at regularly spaced time points $t = 1,2,\ldots,n$. Underlying these observations will be a theoretical stochastic process X_t. A different type of analysis is known as spectral analysis of time series. This is based on the autocorrelation sequence ρ_s of a discrete-time stationary series, x_t, which has a Fourier transformation representation. The spectral representation for x_t can be interpreted as saying that x_t is the sum of an uncountable infinite number of random components, each associated with a particular frequency, and with each pair of components being uncorrelated (Granger & Watson, 1984).

Author Note

Joke H. van Velzen, Research Institute of Child Development and Education, University of Amsterdam. Correspondence concerning this article should be addressed to Joke van Velzen, e-mail: j.h.vanvelzen@sigmetack.com.

Conflict of Interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Citation

Van Velzen, J. H. (2024). Situations and the search for conditional information. *RRREaT: Data-Analytical Techniques, 1*: 01. Retrieved from http://sigmetack.com.

Copyright © 2024