Hutmacher, F. & Wendt, A. N. (Eds.). (forthcoming). *Theory- and model-building in psychology*. Palgrave Macmillan. This is a preprint. Please refer to the publisher's website for the version of record

Toward a More Accurate Notion of Exploratory Research (And Why it Matters)

Uljana Feest (University of Hannover) & Berna Devezer (University of Idaho)

Abstract

The paper analyzes the notion of exploration that can be found in the distinction between exploratory and confirmatory research, which is sometimes appealed to in the metascience literature. We argue that this notion (a) differs in important respects from previous works in exploratory data analysis and (b) contains some counterintuitive assumptions about the nature of exploration. Engaging with works in the history and philosophy of experimentation and modeling, we develop and defend a more comprehensive and accurate notion of exploration and argue that it is better suited for a normative analysis of exploratory research.

Keywords: exploratory research, exploratory modeling, exploratory experimentation, confirmatory/exploratory distinction

1. Introduction

In the debate about the replication crisis of the past 15 years, one focus has been on questionable research practices, such as HARKing (hypothesizing after the results are known)—the practice of re-analyzing experimental data and presenting the results as confirming a hypothesis other than the one researchers started out with (Kerr 1998). It is commonly understood and widely held that even though it is impermissible to treat the results of such a re-analysis as confirmatory of the new hypothesis, such a re-analysis of the data should be accepted in the toolbox of exploratory research. In other words: data can be re-analyzed with the aim of generating new hypotheses, as long as these ideas are not presented as being confirmed by the same data that generated them (Bridges 2022). Within the open science movement, the requirement of preregistration was introduced, in part, to prevent researchers from HARKing and to distinguish clearly between confirmatory and exploratory data analysis. It thereby relies on two assumptions. The first assumption is that preregistration should make it harder for researchers to treat their data as confirming a hypothesis different from the one they set out to test. The second assumption is that “exploratory” data analysis is not a questionable research practice as long as it is properly tagged as such.

Recent work about the philosophy of scientific methodology has suggested that the case against HARKing is not as clear-cut as the above characterization makes it sound (e.g., Devezer et al. 2021, Rubin 2017, Rubin & Donkin 2022, Szollosi & Donkin 2021). Here we take a different approach, however. It is not our aim to discuss whether retrospective data analysis can be confirmatory. Rather, we want to dissect the notion of *exploratory research* as used in the metascience literature. Specifically, we will argue that this notion is (a) too narrow to capture other important ways in which research can be exploratory, and (b) too broad as it does not provide us with normative guidelines to distinguish between epistemically fruitful and epistemically deleterious forms of “exploratory” data analysis. We develop our argument by drawing on the philosophy of science literature about exploratory research (e.g., Steinle 1997, O’Malley 2007, Colaco 2018, Massimi 2019, 2020). This literature, we will argue, provides us with an appreciation of the importance of exploration in both empirical research and theoretical modeling. It also highlights that if we restrict the notion of *exploration* to processes of data analysis, we risk neglecting the processes that led to their creation. This means that there is a danger of treating those data as given rather than debating their quality and relevance.

Section 2 begins by providing an overview of an influential articulation of the exploratory/confirmatory distinction within statistics and psychology offered in the early replication crisis literature (Wagenmakers et al. 2012). We argue that current advocates of exploratory data analysis use the word “exploratory” in a way that diverges not only from a common sense understanding of exploration but also from older treatments of exploratory data analysis (e.g., Tukey 1977). In section 3 we give some background about the practical and experimental turns within philosophy of science in the early 1980s, highlighting that it was before the backdrop of these movements that interest in exploratory research emerged, both within the philosophy of experimentation and the philosophy of modeling. We show that similar impulses drove works stemming from these two traditions, and argue that exploratory and theoretical work need not be mutually exclusive. Section 4 discusses how *exploratory modeling* was highlighted as a specific mode of exploratory research in philosophy of modeling (e.g., Fischer, Gelfert & Steinle 2019) and identifies a particular strand of recent metascience work on exploratory modeling (e.g., Devezer & Buzbas 2023; Guest & Martin 2021) as exemplifying this mode. While section 4 focuses on the systematic variation of modeling or experimental parameters as a crucial ingredient of exploratory research, section 5 expounds on another aspect of exploratory research, namely that it takes place in a situation of conceptual openness. We point to the vagueness of psychological concepts (Meehl 1978) and draw connections to recent philosophical analyses of the relationship between exploratory research and conceptual development (e.g., Feest 2010, Haueis 2023). Section 6 returns to examining whether the distinction between confirmatory and exploratory data analysis has the necessary precision to distinguish between productive exploratory work and questionable research practices. We argue that it does not.

2. On the Exploratory/Confirmatory Distinction in Statistical Data Processing

In the aftermath of the replication crisis, a growing literature has focused on distinguishing between confirmatory and exploratory research. Originally restricted to the null hypothesis significance testing (NHST) framework (despite subsequently being extended to other modes of inference), this dichotomy suggests that exploratory and confirmatory research primarily and firmly diverge in their analytical approaches. Confirmation is said to be characterized by the prior specification of the statistical analyses the researchers intend to carry out, which also means locking in the hypotheses to be tested in an NHST framework, whereas exploration can be performed via post hoc data analyses that Wagenmakers et al. (2012) refer to as “wonky”: “If you carry out a hypothesis test on the very data that inspired that test in the first place then the statistics are invalid (or “wonky”)” (633). Such wonky analyses may overfit the observed data, cherry-pick the results, or misrepresent the procedures implemented, inflating the support for the hypothesis. Such practices, the authors argue, are illegitimate if they are treated as confirmatory of the hypothesis, but are legitimate so long as they are used as part of a process of exploration, not confirmation.[1]

The literature discussed above suggests that researchers should be required to preregister their hypotheses and statistical analyses to prevent confirmatory value from being claimed for an analysis that is “merely” exploratory. Preregistration policies have a lot of prima facie plausibility and we certainly endorse the underlying commitment to transparency and accountability. However, we are skeptical of how this policy has come to be tied to the distinction between confirmatory and exploratory research since those two concepts are less than clear. Here we set aside issues surrounding the definition of “confirmatory analysis” within Wagenmakers et al.’s framework. Instead, we focus on the features they attribute to *exploratory data analysis* to probe underlying assumptions about the nature of exploration.

First, Wagenmakers et al.’s definition of “exploration” strictly relies on the absence of a hypothesis specified before seeing the data. What the authors further imply is that within the context of exploratory research hypotheses would need to be generated based on data and data alone. In this framework, exploration informed by a hypothesis or theory is not explicitly considered a possibility. At any rate, the authors do not specify what a theory-informed exploration might look like. While using the broader label of “exploratory research”, this framework is narrowly limited to exploratory statistical analysis of experimental data rather than providing a comprehensive perspective on exploration in science. Perhaps this statistical data analysis-centered view of exploration invites a comparison to an earlier framework for Exploratory Data Analysis (EDA) envisioned by Tukey (1980, 1993). Tukey and Wagenmakers et al. have in common an exclusive focus on data analytical procedures and agree that for

[1] This asymmetry between what is allowed in exploratory (vs. confirmatory) research is in line with the common perception that exploratory research is inherently less rigorous than confirmatory research. As will become clear in this article, we believe this common perception to be mistaken.

confirmatory analyses, data collection or analysis needs to follow a clear formulation of hypotheses. Tukey's view differs fundamentally in its objective and approach to exploratory analyses, however. Tukey's EDA aims to find questions not necessarily with an eye to generate new hypotheses but to inform how to conduct and analyze future experiments. Instead of recommending haphazardly performing wonky inferential statistical procedures, Tukey emphasizes a thorough numerical or graphical examination of the data without using probability in general or significance tests in particular. He recommends flexible techniques to describe structures that emerge in the data and cautions the analyst to not mistake these structures for natural laws.

Second, in Wagenmakers et al.'s dichotomous framework, performing invalid statistical inference is suggested to be an acceptable, if not recommended, scientific practice as long as it is correctly labeled as exploratory. We argue that even though it is conceivable that good hypotheses may be generated this way, it remains unclear why this method should work or how successful it would be. From a theoretical perspective, there is no reason to believe that misusing or abusing statistical procedures could reliably lead to scientific discoveries (Devezer et al. 2021). The view of exploratory practices as purely data-driven also does not quite fit within a common-sense understanding of *exploration*. For example, we would not send a team of explorers to the Arctic without a plan or reliable equipment and supplies. Why, then, would we send scientists to explore empirical and theoretical spaces unprepared and ill-informed? Moreover, it is not clear how we would distinguish between more or less promising exploration in Wagenmakers et al.'s framework. Good (1983), in his attempt at a philosophy of exploratory data analysis, suggests that a pattern found in EDA should be considered salient or interesting only if it has "a non-negligible subjective or logical probability of being potentially explicable, at least in part" (p. 290). He approaches exploratory analyses with a probabilistic mindset to support valid inferences from the data[2]. In the same spirit, coming out of statistics, there are procedures tailored toward exploratory data analysis which nonetheless return statistically valid results aimed at delineating successful exploratory approaches (e.g., Gelman 2004, Goeman and Solari 2011). As such, exploration via invalid statistical inference remains contested and appears to be insufficient to capture the essence of the process of data exploration.

Third, in Wagenmakers et al.'s framework, selective reporting of desired results is legitimized. This is a particularly puzzling premise given the basic definition of "exploration" which typically contains some reference to a thorough investigation with the aim of learning about something

[2] Notably I. J. Good is a Bayesian statistician and his perspective on EDA is consistent with Cromwell's rule coined by another Bayesian statistician, Dennis Lindley—a rule that reminds us of the importance of maintaining a sense of doubt in our reasoning. Good goes beyond the categorical framing of Cromwell's rule, however, to highlight that we should not readily accept an explanation emerging in exploratory data analysis if we attach to it a vanishingly small prior probability. On the flipside, this would also suggest that if we start exploration with too strong an idea of how to explain the patterns, we would not be able to change our minds in light of the data.

unknown. For example, Tukey (1993) emphasizes that EDA should not exclusively focus on finding happenstance phenomena but also needs to describe some “continuing reality” (p. 7). Tukey appears to suggest that instead of singling out surprising patterns in the data, exploration should also remark on recurring, baseline effects as well as negative or statistically nonsignificant patterns that could be important. Similarly, Höfler et al. (2021) suggest that a valuable exploration should be honest, transparent, and full instead of selectively analyzing the data or cherry-picking desirable results. Moreover, it seems that despite the commitment to a non-theory-driven approach to exploratory data analysis, its selective nature is likely to import biases that will remain unchecked if it is not or cannot be modeled properly.

In light of the three points just mentioned, it seems that the usage of the term “exploratory” in writings such as Wagenmakers et al. pulls in two different directions, each of which is incompatible with plausible (common-sense and statistical) assumptions about exploration. On the one hand, they treat exploratory research as happening in the absence of prior hypotheses and, thus, as not being guided by any theoretical or conceptual assumptions. This goes counter to a common-sense notion of *exploration* as involving at least some basic preliminary understanding of the terrain that is being explored and the tools needed to explore it. On the other hand, they suggest that exploratory research does not need to engage in a systematic investigation of the entire terrain but can, rather, focus on some regularities that were generated by a rather arbitrary choice of procedures. This goes counter to the idea that exploration should be thorough and systematic. If, as we argue, exploration cannot be entirely free of preexisting theoretical and conceptual assumptions, it seems important to us that these assumptions be made as transparent as possible and that their theoretical warrant is discussed. When this doesn’t happen, the worry is that the exploratory process encourages specific conceptual biases to creep in through the back door.

In sum, Wagenmakers et al. (and the metascience literature following their lead) does not embrace the intuitive idea that exploration ought to be guided by some normative framework determining what constitutes good exploration. Nor does it follow prior works on exploratory data analysis (Tukey, Good). Instead, it construes exploration as aiming for serendipitous discoveries, free from scientific or statistical constraints. We see the danger that such a liberal and unprincipled approach is likely to yield invalid statistical inference or spurious results. By contrast, we construe exploration as aiming to provide a carefully planned, systematic mapping of unknown spaces. But this still leaves open the question of how to characterize such a carefully planned and systematic mapping. If we understand exploration to rely on prior assumptions, how should the relationship between those assumptions and the exploratory activity be construed? Can exploration be guided by hypotheses or theories? Does exploration need to be understood in contrast with confirmation? Is there exploratory research beyond data analysis and what does it look like? How flexible should exploratory research be? What criteria characterize good scientific exploration? In the following, we turn to the literature on exploratory research coming out of philosophy of science. Section 3 introduces the reader to the context in which a focus on exploratory research emerged within philosophy of science. Section 4 explicates some functions of exploratory modeling.

3. The Emergence of Interest in Exploratory Research in Philosophy of Science

Within the field of the history and philosophy of science (HPS) exploratory research was not an object of analysis until the late 1990s when two papers appeared the same year that each used the term "exploratory experiment" (Burian 1997; Steinle 1997). These authors identified what they took to be a specific function of experiments that had not been covered by previous scholarship. The emergence of philosophical interest in exploratory experimentation has to be placed within the general context of two movements that emerged within HPS and STS that are sometimes referred to as the "practice turn" and the "new experimentalism," respectively. In general terms, all of these approaches are characterized by a rejection of the distinction between the contexts of discovery and justification.[3]

The practice turn in philosophy of science was informed by several important insights. First, practices take place in space and time, they are situated and temporal. Looking at scientific practice, thus, decentered (at least for some scholars) the traditional interest in a purely logical relationship between theory and evidence in favor of a more historically informed study of knowledge production. In addition, it drew attention to material aspects of science, such as experimentation, the use of scientific instruments, and practices of modeling. One feature shared by many instances of this turn to practice was the recognition of an epistemically important "middle realm" that does not easily fit in the traditional dichotomy between theory and evidence.

The spirit of experimentalism, as it emerged in the 1980s, is epitomized in Ian Hacking's famous statement that "experimentation has a life of its own" (Hacking 1983, 150), by which he meant to highlight that our philosophical interest in scientific experiments should not be restricted to their function in theory testing. Subsequently, one prominent alternative proposal concerning the function of experimentation was articulated with the concept of *exploratory experimentation* already mentioned above. For example, Friedrich Steinle (1997) took this concept to describe a

[3] This distinction has traditionally often been understood as implying that the rationality of science plays out in the logic of theory-testing only, whereas processes of theory-generation are not - and need not be - governed or constrained by rational method and therefore do not fall within the purview of philosophy of science. While this reading of the distinction seems to be at the heart of the intuitive appeal of the exploratory/confirmatory dichotomy, there is a long-standing consensus within philosophy of science that it is problematic, though there are several different construals of what makes it problematic (Hoyningen-Huene 1987). For our purposes, it suffices to say that normative analyses of knowledge generating activities have become an established part of contemporary philosophy of science.

practice of systematic variation of experimental parameters. A second (related) reading of the importance of exploratory experimentation is that sometimes experiments about a given subject matter take place in the absence of theory and, thus, in a context of conceptual openness (Steinle 1997, Burian 1997). Other scholars also pointed to the poverty of existing philosophical accounts of the purpose of scientific data production. Bogen and Woodward (1988), in particular, argued against the common assumption that theories predict observations, claiming instead that theories predict phenomena, which are not straightforwardly observable. With this, they pointed to the local and idiosyncratic conditions, under which experimental evidence for phenomena is produced. In addition, Woodward (1989) also argued that one of the functions of experimental research can be to establish the existence of specific phenomena in a given domain, as opposed to testing theories about them. In turn, Feest (2010) has argued that this kind of activity requires (at least) rudimentary concepts of the phenomena, which inform the ways in which tests about them are operationalized. Importantly, this tradition of thinking about exploratory aspects of research focuses, among other things, on the *production* of data. This is in stark contrast with the literature about exploratory data analysis we discussed in the previous section, which starts off with *already existing* data.

The practice turn was not restricted to experimental practices, however, but also came with a rising interest in scientific *modeling* as a key feature of scientific practice. Just like Ian Hacking had resisted an understanding of science as aiming at representation (highlighting practices of experimental intervention instead), philosophers interested in modeling also moved away from thinking of models as mere representations toward an understanding of models as tools of exploration (Gelfert 2016). This way of thinking was spearheaded by Morgan & Morrison (1999), who coined the expression "models as mediators," thereby capturing both the idea of models as instantiating an existing theory (e.g., van Fraassen 1980) and the idea of modeling as being a form of theorizing, where the latter is construed as a process by which our scientific understanding of a specific subject matter is pushed forward (e.g., Bailer-Jones 2000). Importantly, on this understanding, theorizing about a given subject matter is not mutually exclusive with its exploration. In other words, models can be seen as allowing researchers to investigate features of their objects of investigation while at the same time engaging in a form of theoretical work about these objects.

When we look at the history of the philosophy of experimentation, it becomes clear that here, too, there was a gradual shift away from thinking of exploration only in connection with a specific type of experiment toward a more expansive understanding of exploration as a type of activity that integrates theoretical and experimental components. This development happened, in part, as philosophers of science engaged with Steinle's initial juxtaposition of exploratory experimentation and theory-driven experimentation: "[T]heory-driven experiments are typically done with quite specific expectations of the various possible outcomes in mind," whereas "[e]xploratory experimentation … is driven by the elementary desire to obtain empirical regularities…" (Steinle 1997, S70). This dichotomy was challenged by subsequent works that pointed out that some forms of experiments are aptly described as "exploratory" even though they clearly rely on a lot of theoretical assumptions, including both auxiliaries and assumptions about the very subject matter under exploration (see for example Colaço 2018; see also Feest & Steinle 2016).

If we follow this line of reasoning, it seems that exploratory experimentation can involve various degrees of theory, raising the question of how clear-cut the distinction between theory-driven and exploratory experiments really is. It is for this reason that we adopt the term "exploratory research," which we take to refer not to a specific type of experiment but to an ongoing, iterative process of trying to establish empirical regularities and mechanisms that are connected to a given object of research, to delineate the boundaries of that object, and to provide theoretical descriptions of them (Feest 2025, chapter 4). In this regard, we follow O'Malley (2007), who has argued that the very distinction between exploratory experimentation and theory-driven experimentation may be an oversimplification of real investigative processes, particularly in highly complex domains that increasingly use massive amounts of data. Accordingly, we wish to retain the notion of "exploratory" to do justice to Steinle's insight about the importance of varying contexts as part of the dynamics of systematic explorations. Importantly, however, we argue that the systematic variation of parameters need not be restricted to the variation of *experimental* parameters. Mathematical modeling and computer simulations can play a very similar role and, thus, should be included in the repertoire of tools used in exploratory research.

4. On the role of mathematical models in exploratory research

The upshot of the above is that we retain Steinle's (1997) insight that exploratory research involves the systematic variation of parameters while following more recent work in philosophy of science in rejecting the idea that such a research activity is (or should be) theory-free. In this section, we argue that this way of thinking about exploratory research dovetails with recent work on exploratory modeling, which has also argued that modeling can be exploratory despite being closely tied in with theoretical work.

The term "model" is notoriously hard to pin down as it refers to many kinds of things, ranging from physical objects to fictional and abstract objects to set-theoretic structures and equations (cf. Frigg & Hartmann 2024). There is widespread agreement, however, that models and modeling can serve a range of epistemic functions, such as prediction, explanation, guiding data collection, discovering new questions, bounding outcomes to plausible ranges, illuminating uncertainties, demonstrating tradeoffs, and challenging robustness of scientific theories by introducing perturbations (Epstein 2008). Many of these functions involve some level of exploration in the theoretical problem space, in the model space, or in the data.

To understand what this means, we highlight a key distinguishing characteristic of model-based approaches in science, namely that they rely on mappings. These mappings can be physical, verbal, formal, and so on. It is by virtue of such mappings that modeling can help researchers explore, understand, and explain complex unknown phenomena (e.g., mind and behavior, see van Rooij, 2022) by studying a simpler model system that resembles the target system in certain desired aspects. Modeling may, thus, allow researchers to engage in indirect exploration of a target system, even when they set out to test theoretically informed hypotheses about the behavior of the model system. For example, we may set up an in-silico experiment to test whether a formal model behaves in a certain way because of a specific assumption and manipulate that assumption via simulations. As a result, we will be able to make a decision about the sensitivity of model outcomes to that assumption, but we won't have tested whether that assumption holds for the target system. We will, however, have explored the potential implications of certain conditions for the target system.

Clearly, the modeling activities described in the previous paragraph are "theoretical" in some sense, insofar as they rely (minimally) on decisions about how to represent the target system. In addition, the existing literature about formal modeling offers a range of proposals as to how exploratory modeling can play out for different target systems in ways that contribute to theorizing. For example, when the target system is a theoretical framework, formal modeling invites exploration into the nature of theories by enforcing an explication of their assumptions and implications (Smaldino 2017). Particularly since formal models need to be expressed in mathematical language or computational code, creating a full and transparent specification of the knowns and unknowns of the theory becomes a necessity. Suppes (1962) has argued that even in the absence of sufficient theoretical development, models of experiment and models of data can be used to guide scientific inference. This is relevant for our purposes because in psychology it is common to perform experiments with the aim of testing specific hypotheses in the absence of well-specified theories (Feest 2025). Guest and Martin (2021) suggest that in these situations formal modeling may induce open theorizing and help make opaque ideas underlying the hypothesis more transparent and explicit.

Suppes (1962) provides a broader analytical framework by conceptualizing research as involving a hierarchy of models. He defines *models of data* as concerning all aspects of the experiment that can be used in statistical tests aimed at testing the adequacy of theories such as homogeneity and invariance of effects. By contrast, he argues that *models of experiment* exist at a higher level in the model hierarchy and involve the choice of experimental parameters such as sample size. Crucially, Suppes believes that when theoretical knowledge consists of a collection of heuristics rather than a well-specified logical structure, models of experiment and data could still be employed to guide empirical research approaches. In this way, his account captures our contention that formal modeling can be exploratory. In the kind of model hierarchy proposed by Suppes, it's not necessarily the nature of modeling activity that differs between levels so much as the target system realized by the model, whether it be a theory, an experiment, or observations about a phenomenon. Regardless of the level models operate in, specifying what they are helps researchers articulate the logical arguments they are making about the relationship among parts within a target system, lay out all assumptions, and examine the appropriateness and implications of these assumptions. This counterfactual ability to infer conclusions that follow from given assumptions, allows modelers to revise their models by juxtaposing them with reality. Such an iterative process aims to move between different levels of the model hierarchy by rigorous triangulating, testing, and updating (Devezer and Buzbas 2023).

This big-picture view of scientific modeling helps us appreciate the role models can play in exploratory research. In particular, the above sketch of formal modeling (a) captures the iterative character of exploration and (b) highlights the role of counterfactual reasoning in the exploratory process. Pursuing a similar argument, Massimi (2019) has argued that "targetless" and "hypothetical" models can help us generate modal knowledge about what is causally or objectively possible, "thanks to their specific kind of imagining or conceiving nonveridical scenarios" (Massimi 2019, 871). In computational approaches, this is often achieved via performing simulations under a formal model and observing the sensitivity of the model to changes in initial conditions (Gelfert 2016, 73). This type of procedure is also illustrated by Devezer et al. (2019), which advances a mathematical model of the scientific process to study the sensitivity of certain scientific outcomes (such as reproducibility of results and speed of discovery) to changes in parameter values (such as different research strategies, level of system noise, and choice of statistical methods). The authors argue that in-depth exploration of this model under full model specification allows the researchers to observe plausible outcomes under most conservative conditions as well as the boundaries of what is possible under extreme assumptions.

From the above discussion, it follows that modeling can be exploratory in at least four interrelated senses:

1) Theoretical exploration: Modeling can be used to explore the implications of an already fully formalized theory with a clear logical structure. In this sense, it may help generate theoretical predictions or hypotheses to be tested and guide data collection.

2) Conceptual exploration: Modeling may help explicate hazy theoretical assumptions about the subject matter, thereby making them less hazy and allowing for their robustness to be probed by introducing perturbations. In this sense, it aids in the gradual, iterative building of theories.

3) Data exploration: Empirical modeling may help explicate assumptions made in experimental design and explore the relationship between design constraints and the data generated. To this end, model exploration does not depend on any degree of formal theorizing and can be minimally or vaguely informed by some informal theoretical background.

4) Modal exploration: Modeling can be used to explore the boundaries of what is possible and what is inevitable under different theoretical assumptions or initial conditions. In this capacity, it may inform the discovery of new research questions, expose theoretical gaps, and identify plausible bounds for parameter values.

5. Conceptual Openness in Exploratory Research

In the previous section, we argued that model-based simulation as an exploratory activity is akin to exploratory experimentation as it involves varying model (vs. experiment) parameters and observing the implications to map out the outcome space. As we have laid out, both experimentation and modeling can be exploratory even if they (a) presuppose theoretical or conceptual assumptions as auxiliaries or (b) contribute to the development of a theory of the object of exploration.

In this section we turn to a second necessary ingredient of exploration that was briefly touched on above, namely, the idea that exploration involves venturing into an unknown or little-understood space. Within the literature about exploratory experimentation, this has at times been described in terms of the idea of *conceptual openness*. Indeed, for Steinle (1997), a crucial feature of exploratory experimentation was that it could result in a conceptual restructuring of an entire field. Another way of putting this is to say that exploratory research is required when the descriptive and explanatory concepts for a novel domain are not yet well articulated. As laid out in the previous two sections, the notion of *conceptual openness* should not be confused with the idea that researchers start with a blank slate. The idea, rather, is that exploratory research is called for when it is not yet clear to what extent existing concepts are applicable. As a result of exploratory research, then, concepts can get extended and changed in sometimes unanticipated ways.

The notion of exploration as taking place in conceptually open spaces, while necessarily building on existing concepts is also hinted at by Gelfert (2016), who points out that exploration can be specific (targeting a particular object) or nonspecific. Both types of exploration presuppose conceptual assumptions. They only differ in the extent to which they draw on prior assumptions and in the level of specificity of those assumptions. Illustrating this with a common-sense example, we may say that when someone explores an unknown environment in search of food, they will clearly presuppose not only the concept of *food* but also some basic criteria of how to distinguish food from non-food (such as smell and look). This is perfectly compatible with the idea that such exploration can result in the discovery of previously unknown forms of nutrition and extend our very concept of what is edible. Haueis (2023, 2024) has recently begun developing a principled philosophical account of the relationship between exploratory research and conceptual extension in science. Turning specifically to psychology, such an understanding of exploration as conceptual articulation and extension is also hinted at by Hutmacher & Franz (2024), who argue that the inherent vagueness of psychological concepts calls for research that explores the boundaries of their referents.

Within the history and methodology of psychology, we encounter a similar sentiment already in Paul Meehl's (1978) pronouncement that the scientific concepts of psychology are "open." According to Meehl, this openness arises as a result of "the indefinite extensibility of our provisional list of operational indicators, …. the empirical fact that indicators are only probabilistically, rather than nomologically, linked to the inferred theoretical construct, and the fact that most of our theoretical entities are introduced by an implicit or contextual definition" (Meehl 1978, 815). This statement acknowledges that psychological concepts are laden with pre-theoretical content, which needs to be both exploited and explored (see also Feest 2025). As Meehl points out, psychological concepts are rarely fully explicated, prompting him to call on psychologists to "learn to be sophisticated and rigorous in their metathinking about open concepts at the substantive level" (815) and to embrace the possibility of "conceptual drift" (816). Meehl envisions such conceptual drift to occur in close coordination with "psychometric drift" (816), by which he means the principled improvement of measurement tools. This is important because it reveals an understanding of the exploratory process as involving an iterative process of improving measurement tools (and, thus, empirical data) and developing one's concepts. We argue that this vision of the investigative process in psychology is highly compatible with the account of exploratory research we have sketched in the previous sections. Recall, for example, that one of the rationales for exploratory modeling was that such modeling forces researchers to make the hazy content of their concepts explicit, thereby allowing for their investigation by way of systematic variation of model parameters. Likewise, Meehl's emphasis on the close connection between psychometric and conceptual progress also points to the need for a notion of exploratory research that does not begin after the process of data generation (in the service of a specific hypothesis-test) has been completed but begins with the very design of the experiment mode, or test that enable the generation of data.

Summing up, we endorse a notion of *exploratory research* that highlights two crucial features: the systematic variation of experimental and model parameters on the one hand and the conceptual openness of the research situation on the other. As we have argued, both of these features are compatible with the idea that the exploratory process can be guided and constrained by conceptual or theoretical assumptions as well as with the idea that exploratory research can contribute to theory formation. For example, the systematic variation of parameters (both in experiments and models) can be a tool for theory construction and conceptual development.[4] Moreover, we have pointed out that our analysis is compatible with a range of existing positions in data science, formal modeling, and philosophy of science going back several decades. On this account, it is an essential feature of exploratory research that it unfolds across a variety of research activities and projects over time rather than within the bounds of a single project or research article. In our view exploration can start from theories, vague conceptualizations, apparent paradoxes, observations, and even specific research results and continues until the target object is fully fleshed out and most of the questions surrounding it—whether they're modal, theoretical, conceptual, or empirical—have been raised and examined. In this regard, our account of what is *exploratory* in exploratory research is - in a certain sense - much broader than the one that is implicit in the confirmatory/exploratory distinction (Wagenmaker et al. 2012): It pertains not only to a specific kind of data analysis but to a larger investigative endeavor that can also include theorizing, modeling, experimenting, etc. Moreover, it provides us with diagnostic tools to talk about the standards that high-quality exploratory research should meet. As we will argue in our concluding section, those standards are, ultimately, more demanding than those implicit in Wagenmaker et al.'s account of exploratory data analysis. The remaining question, then, is why our analysis should be preferred.

6. Conclusion: Why it Matters

Per the analysis presented in this paper, exploratory research is not an individual, clearly delineated, precise research activity but a process of mapping a given object of research in a theoretical and empirical space, by iteratively negotiating between theory and data to first make boundaries, unknowns, uncertainties explicit, and then gradually uncover what is knowable. In our analysis, exploratory research can occur both against the backdrop of strong theory and in situations that are characterized by conceptual openness, though (especially in this latter case) it needs to be informed by the aim to do a thorough search of the space of possibilities.

This understanding of exploratory research differs from exploratory data analysis as conceptualized by Wagenmakers et al. (2012) and adopted in the post-replication crisis metascience literature. As we explained above, that literature restricts exploration to statistical analyses that are formulated and performed after having seen experimental data. The only aim

[4] As similar examples of novel forms of exploratory experimentation, we may contrast the notion of metastudy by Baribault et al. (2018), who suggest radically randomizing experimental parameters within a so-called metastudy for the purposes of robust theory testing, with the notion of integrative experimental design by Almaatouq et al. (2024), who suggest mapping the design space of possible experiments associated with a research question to inform iterative theory development.

of such analyses is to generate hypotheses by singling out statistical patterns emerging in the data. Thus, according to the view we reject here, the defining characteristic of exploratory analysis is that it does not test hypotheses determined prior to any knowledge of the data.

Now, the question might be raised whether this is a merely terminological point. Surely, Wagenmakers et al. (2012) and other authors are entitled to use the term "exploratory" any way they want. We have argued in this paper that their notion of exploration is counterintuitive insofar as it does not seem to require background assumptions or systematic search. However, this does not imply that there might not be a perfectly legitimate, more restricted notion of *exploration* that applies to the kind of exploratory data analysis suggested by them. Perhaps our broader notion of exploration is compatible with the more narrow one proposed by them? To argue against this possibility, we need to show that a purely data-driven, non-theoretical understanding of exploratory data analysis is not only counterintuitive but also epistemically counterproductive.

To examine this question, let's pick up the common argument for preregistration that we presented at the outset, namely, that it is a valuable tool (a) to prevent researchers from misrepresenting their post-hoc theorizing as confirming a hypothesis, and (b) to distinguish confirmatory from "merely" exploratory research. Our analysis grants that preregistration can contribute to transparency. However, we are skeptical of the way this argument has been tied to the distinction between confirmatory and exploratory data analysis. For one thing, a hypothesis can be confirmed (or disconfirmed) by the data regardless of whether the analysis was preregistered (provided that it was done by an honest researcher). But here we are interested in a more substantial question, namely whether exploratory data analysis (again, in the sense of Wagenmakers et al.) is either necessary or sufficient for that analysis to be truly exploratory. We argue that it is not.

The non-necessary part would likely be acknowledged by advocates of exploratory data analysis since they don't need to commit to exploratory data analysis being the *only* kind of exploratory research. Once this is granted, however, this calls for a more in-depth analysis of the notion of *exploratory research*. Such an analysis has been provided in this paper. A crucial upshot of our analysis is that exploratory research starts out by acknowledging the conceptual openness of the research situation, while also making use of existing presuppositions and exploring the relevant epistemic space. This is done by varying experimental or model parameters in order to further the theoretical or empirical understanding of a given subject matter. Once this is acknowledged, however, it seems that exploratory data analysis also cannot be sufficient for exploratory research. The data of such an analysis are the outcome of an experiment that was designed to test a specific hypothesis. So, there is a sense in which the data exhibit a path dependence on that previous hypothesis. But there is no guarantee as to the quality of that hypothesis, and hence it is not clear that the data are of any relevance to the subject matter. This means that the to-be-explored space is potentially already severely hampered by the availability of only a limited set of data. This runs counter to a feature we have identified as important to exploration, namely the aim to map the space of possibilities. If, on the other hand, researchers have conceptual or theoretical reasons for believing the data to be of interest, this reveals that they are implicitly acknowledging our thesis that exploratory research is an iterative and ongoing process that builds on and expands existing conceptual, theoretical, and instrumental tools. In that case, we argue, exploratory research requires that those tools be made explicit so that they can be analyzed, discussed, and probed more systematically.

Summing up, in this paper we have drawn on some relevant philosophy of science literature to argue that the notion of exploration implicit in the dichotomy between confirmatory and exploratory research is quite impoverished, and therefore insufficient for fully capturing the nature of scientific exploration. By this, we mean not merely that it only covers a part of the exploratory process (i.e., the data analysis). Rather, we argue that if our understanding of exploration is adopted, this reveals either (a) that exploratory data analysis, on its own, does not meet the criteria for exploratory research at all (even if, by chance, it occasionally happens upon an interesting hypothesis), or (b) that whatever makes such analyses exploratory (e.g., drawing on existing conceptual and technological tools with the aim of investigating possibility spaces through modeling or experimenting) needs to be brought out into the open. This is the sense in which we fully embrace the open science movement's call for transparency (see also Leonelli 2023 for an account of open science that we are sympathetic to).

REFERENCES

Almaatouq, Abdullah; Griffiths, Thomas; Suchow, Jordan; Whiting, Mark; Evans, James; Watts, Dunkin (2024). Beyond playing 20 questions with nature: Integrative experiment design in the social and behavioral sciences. *Behavioral and Brain Sciences* 47, e33: 1–70. doi:10.1017/S0140525X22002874

Bailer-Jones (2000). Modelling Extended Extragalactic Radio Sources. *Stud. Hist. Phil. Mod. Phys.* 31, No. 1, pp. 49-74

Baribault, Beth, Chris Donkin, Daniel R. Little, Jennifer S. Trueblood, Zita Oravecz, Don Van Ravenzwaaij, Corey N. White, Paul De Boeck, and Joachim Vandekerckhove. (2018). Metastudies for robust tests of theory. *Proceedings of the National Academy of Sciences,* 115, no. 11: 2607-2612.

Bridges, A.J. (2022). Hypothesizing After Results Are Known: HARKing. In: O'Donohue, W., Masuda, A., Lilienfeld, S. (eds) *Avoiding Questionable Research Practices in Applied Psychology*. Springer, Cham. https://doi.org/10.1007/978-3-031-04968-2_8

Colaço, David (2018). Rethinking the Role of Theory in Exploratory Experimentation. *Biology and Philosophy* 33 (5–6):article 38. https://doi.org/10.1007/s10539-018-9648-9.

Devezer, Berna & Buzbas, Erkan O. (2023). Rigorous Exploration in a Model-Centric Science via Epistemic Iteration. *Journal of Applied Research in Memory and Cognition*. Vol. 12, No. 2, 189–194 ISSN: 2211-3681 https://doi.org/10.1037/mac0000121

Devezer, Berna; Nardin, Luis Gustavo; Baumgaertner, Bert; Buzbas, Erkan O. (2019). Scientific discovery in a model-centric framework: Reproducibility, innovation, and epistemic diversity. *PLoS ONE* 14(5): e0216125. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0216125

Devezer, Berna; Navarro, Danielle J.; Vandekerckhove, Joachim; Buzbas, Erkan O. (2021). The case for formal methodology in scientific reform. *Royal Society Open Science.* 8200805 http://doi.org/10.1098/rsos.200805

Epstein, Joshua M. (2008). Why Model?. *Journal of Artificial Societies and Social Simulation* 11(4)12 <https://www.jasss.org/11/4/12.html>.

Feest, Uljana (2025). *Operationism in Psychology. An Epistemology of Exploration*. University of Chicago Press

Frigg, Roman and Stephan Hartmann (2024). Models in Science. *The Stanford Encyclopedia of Philosophy* (Fall 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2024/entries/models-science/>.

Gelfert, Axel (2016). *How to do science with models. A philosophical primer*. Springer briefs in philosophy. Springer-Verlag.

Gelman, Andrew. (2004). Exploratory Data Analysis for Complex Models. *Journal of Computational and Graphical Statistics* 13: 755–779. https://doi.org/10.1198/106186004X11435.

Goeman, Jelle J., and Aldo Solari. (2011). Multiple Testing for Exploratory Research. *Statistical Science* 26. https://doi.org/10.1214/11-STS356.

Good, I. J. (1983). The Philosophy of Exploratory Data Analysis. *Philosophy of Science* 50: 283–295.

Guest, Olivia, and Andrea E. Martin. (2021). "How computational modeling can force theory building in psychological science." *Perspectives on Psychological Science* 16, no. 4: 789-802.

Hacking, Ian. (1983). *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge University Press, Cambridge.

Haueis, Philipp (2023). Exploratory Concept Formation and Tool Development in Neuroscience. *Philosophy of Science*. 90(2): 354-375. doi:10.1017/psa.2022.79

Haueis, Philipp (2024). A Generalized Patchwork Approach to Scientific Concepts. *The British Journal for the Philosophy of Science* 75 (3): 537-795

Hoyningen-Huene, Paul (1987). Context of Discovery and Context of Justification.. Studies in History and Philosophy of Science 18(4), 101-515

Höfler, M., Scherbaum, S., Kanske, P., & Miller, R. (2021, April 4). A call to accompany pre-registration with means to valuable and effective exploration. https://doi.org/10.31234/osf.io/psfqw

Hutmacher, Fabian & Franz, David J. (2024). Approaching psychology's current crises by exploring the vagueness of psychological concepts: Recommendations for advancing the discipline. *American Psychologist.* Advance online publication. https://doi.org/10.1037/amp0001300

Kerr, Norbert (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review* 2(3), 196-217.

Leonelli, Sabina (2023). *Philosophy of Open Science*. Cambridge University Press

Massimi, Michela (2018). Perspectival modeling. *Philosophy of Science* 86 (5):869-881.

Massimi, Michela (2019). Two Kinds of Exploratory Models. *Philosophy of Science* 86, 869–881.

Meehl, Paul (1978). Theoretical Risks and Tabular Asterisks: Sir Karl, Sir Ronald, and the Slow Progress of Soft Psychology. *Journal of Consulting and Clinical Psychology* 46, 806-834

O'Malley, Maureen (2007). Exploratory Experimentation and Scientific Practice: Metagenomics and the Proteorhodopsin Case. *History and Philosophy of the Life Sciences* 29 (3): 335–58.

Rubin, M. (2017). When Does HARKing Hurt? Identifying When Different Types of Undisclosed Post Hoc Hypothesizing Harm Scientific Progress. *Review of General Psychology*, *21*(4), 308-320.

Rubin, Mark, & Donkin, Chris (2022). Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests. *Philosophical Psychology* 1–29. https://doi.org/10.1080/09515089.2022.2113771

Smaldino, Paul E. (2017). Models are stupid, and we need more of them. In Robin Vallacher, Stephen Read, Andrzej Nowak (eds.), *Computational social psychology*, pp. 311-331. Routledge.

Steinle, Friedrich. (1997). Entering New Fields: Exploratory Uses of Experimentation. *Philosophy of Science* 64: 65 – 74

Suppes, Patrick (2009). Models of data. In Ernest Nagel, Patrick Suppes & Alfred Tarski (eds.), Provability, Computability and Reflection. Stanford, CA, USA: Elsevier.

Szollosi, Aba and Donkin, Chris (2021). Arrested Theory Development: The Misguided Distinction Between Exploratory and Confirmatory Research. Perspectives on Psychological Science 2021, Vol. 16(4) 717–724

Tukey, John W (1977). *Exploratory Data Analysis. Addison-Wesley Publishing Company*

Tukey, John W. (1980). We Need Both Exploratory and Confirmatory. *The American Statistician*, Vol. 34, No. 1, pp. 23-25

Tukey, John W. (1993). Exploratory data analysis: past, present and future. Tech. rep., PRINCETON UNIV NJ DEPT OF STATISTICS.

van Rooij, Iris (2022). Psychological models and their distractors. *Nature Reviews Psychology,* 1: 3, 127-128.

Van Fraassen, Bas (1980). *The Scientific Image*. Oxford University Press

Wagenmakers, Eric-Jan; Wetzels, Ruud; Borsboom, Denny (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science* 7(6) 632–638