SportRχiv

Part of the Society for Transparency, Openness and Replication in Kinesiology (STORK)

Preprint
not peer reviewed

Exploratory Research in Sport and Exercise Science

For correspondence:
massimiliano.ditroilo@ucd.ie

Massimiliano Ditroilo[1], Cristian Mesquida[2], Grant Abt[3], Daniël Lakens[2].

[1] School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

[2] Human Technology Interaction, Eindhoven University of Technology, Eindhoven, The Netherlands

[3] School of Sport, Exercise and Rehabilitation Sciences, Faculty of Health Sciences, The University of Hull, UK

Please cite as: Ditroilo, M., Mesquida, C., Abt, G., & Lakens, D. (2024). Exploratory Research in Sport and Exercise Science. *SportRχiv*.

ABSTRACT

Quantitative exploratory research implies a flexible examination of a dataset with the purpose of finding patterns, associations, and interactions between variables to help formulate a hypothesis, which should then be severely tested in a future confirmatory study. In many fields, including sport and exercise science, exploratory research is not openly reported. At the same time, experts agree that most of the research we conduct is indeed exploratory, and that exploration is a crucial step in scientific knowledge generation. Using a flowchart, we review how data are typically collected and used, and we distinguish exploratory from confirmatory studies by arguing that data-driven analyses, where the Type I and Type II error rate cannot be controlled, is what characterises exploratory research. Even if a study tests a hypothesis in an error-controlled manner, often it also includes exploratory analyses on the data. We ask which factors increase the quality and value of exploratory analyses, and highlight large sample sizes,

All authors have read and approved this version of the manuscript. This article was last modified on Month 09, 2024.

Authors Abt @*grantabt* and Lakens @*Lakens* can be reached on X (Twitter).

uncommon sample compositions, rigorous data collection, widely used measures, observing a logical and coherent pattern across multiple variables, and the potential for generating new research questions as the main factors. Finally, we provide guidelines for carrying out and transparently writing up an exploratory study.

INTRODUCTION

Due to their practical experience, direct contact with sports players, and the ability to collect large amounts of data, strength and conditioning coaches and performance analysts often collaborate with academia. It is not uncommon for them to complete a PhD based on the data they have available from the athletes that they work with. Similarly, a sports and exercise scientist may have good connections with the sport industry to provide consultancy services to teams and individual athletes and get access to data related to their research topic. This is also the case for some of the authors of this article. On multiple occasions we have been approached by the strength and conditioning coach or the performance analyst of a team with a proposal, which approximately sounded like this: “We have a lot of data from athletes we have tested in the last two years, we should try and get a scientific paper out of it. Are you interested?” If we said yes, then the next step would have been to explore the data through several unplanned analyses, such as computing the correlation between variables, testing for body bilateral asymmetries, or measuring seasonal variations in performance.

The goal of these analyses is often to find statistically significant results of hypotheses tests. Showing an effect (or perhaps showing no effect, if it is an eye-catching result) would most likely convince the editor and the reviewers that the study is worth publishing. Even though in our experience many research projects start by exploring a rich and interesting dataset, it is incredibly rare to see researchers reporting this transparently in their manuscripts. Instead, journal publications are presented as hypothesis testing, as if this is the only worthwhile type of scientific article to write. Not only do researchers mostly hide the exploratory nature of their investigation behind a veneer of ‘confirmatory’ tests, but they also limit the potential that a full and transparent description of the exploratory analyses could have had for the scientific community. An exploratory study might add more value in terms of the ideas it will allow others to generate, and the usefulness of statistics that are reported for follow-up research, than presenting the same study as confirmatory.

In this article, we will discuss the core differences between quantitative exploratory and confirmatory research. We will also make the case that sport and exercise science (among other disciplines) would benefit from much more transparently reported exploratory research. Finally, we provide advice on what makes for a better exploratory study, and how exploratory research should be conducted and reported.

EXPLORATORY VS. CONFIRMATORY RESEARCH

Exploratory research involves the free and flexible examination of a dataset with the goal of detecting patterns and associations that may eventually lead to the generation of new hypotheses (Höfler et al., 2022; Kimmelman et al., 2014). In contrast, confirmatory research attempts to test a prespecified hypothesis that may or may not be confirmed by the data (Wagenmakers et al., 2012). Even though these are established descriptions of exploratory and confirmatory research, in this paper we propose narrower definitions where hypothesis tests are confirmatory when their error rates (i.e., Type I and Type II) are controlled, and exploratory when the error rates are not controlled. We will develop the arguments for this more precise definition in the next sections. In most research lines exploratory research necessarily precedes, and inspires, confirmatory research. As Tukey (1980) wrote: "Ideas come from previous exploration more often than from lightning strokes". More recently a growing number of scientists are making their voice heard about the importance of exploratory research (Höfler et al., 2022; Kimmelman et al., 2014; Scheel et al., 2021; Schwab & Held, 2020; Tong, 2019). However, even though the goal of the two types of research is well established (e.g. Schwab & Held, 2020), exploratory research is sometimes regarded as a second-league type of research. This leads to the practice of 'intransparent exploration', (i.e. exploratory studies presented as if they were confirmatory, Höfler et al., 2022).

Exploratory research is well established in other disciplines, such as clinical studies, which is reflected in the distinct phases of clinical trials (ICH, 1998, 2022), where early phases focus more on exploration, and only phase 3 clinical trials perform a confirmatory test of the treatment. Beyond clinical research, a recent paper suggested implementing a distinction between exploratory, confirmatory, and generalizability studies in the field of animal studies trialling disease therapies (Mogil & Macleod, 2017). In other fields such as ecology and consumer psychology the exploratory nature of research might not be explicitly acknowledged, even though most research has been argued to be exploratory (Nilsen et al., 2020; Pham & Oh, 2021). Researchers in brain stimulation have called for more exploratory studies to identify the sources of variability in response to different stimulation techniques (Hussain & Cohen, 2017). In economics the value of this type of research has been highlighted, especially because of budget constraints when it is challenging to perform confirmatory research with controlled error rates (Olken, 2015).

EXPLORATORY RESEARCH IN SPORT AND EXERCISE SCIENCE

It appears that in sport and exercise science a large proportion of research is exploratory in nature simply because the main goal of data collection is often to monitor athletes during training or competition. These datasets can subsequently be analysed to discover patterns and associations, even if they were not explicitly designed with any research question in mind. At the same time, explicitly acknowledged exploratory research is rare in the sports and exercise science literature. To examine this, we searched Scopus for papers published in the last 5 years, using the word ‘sport’ or ‘exercise’ in either title, abstract, or keywords (n = 296,045), and the word ‘exploratory’ appeared in 1.6% (n = 4,737) of those papers (search conducted on 23rd April 2024). From this initial search, we then extracted all the quantitative research papers displaying the word 'exploratory’ in the title and published in journals listed in the category “Sport Sciences” in the Clarivate Journal Citation Reports (n=16). We found that the word ‘exploratory’ was transparently used in only one paper (according to the original definition we have described in section ‘Exploratory vs. confirmatory research’). This study had a very large sample size and reported the exploration of clusters and predictors. The other 15 studies were either intervention studies (n=9), four of which tested at least one hypothesis, or descriptive/observational studies examining differences, associations, or predictions (n=6). All these papers made claims, recommendations, and conclusions. The data seem to confirm our experience that exploratory research is on one hand labelled incorrectly, and on the other hand not commonly transparently undertaken in sport and exercise science.

HOW TO NAVIGATE BETWEEN EXPLORATORY AND CONFIRMATORY RESEARCH

The flowchart presented in Figure 1 is an attempt, with some inevitable simplifications, to provide an overview of how exploratory and confirmatory studies can be differentiated in sport and exercise science.

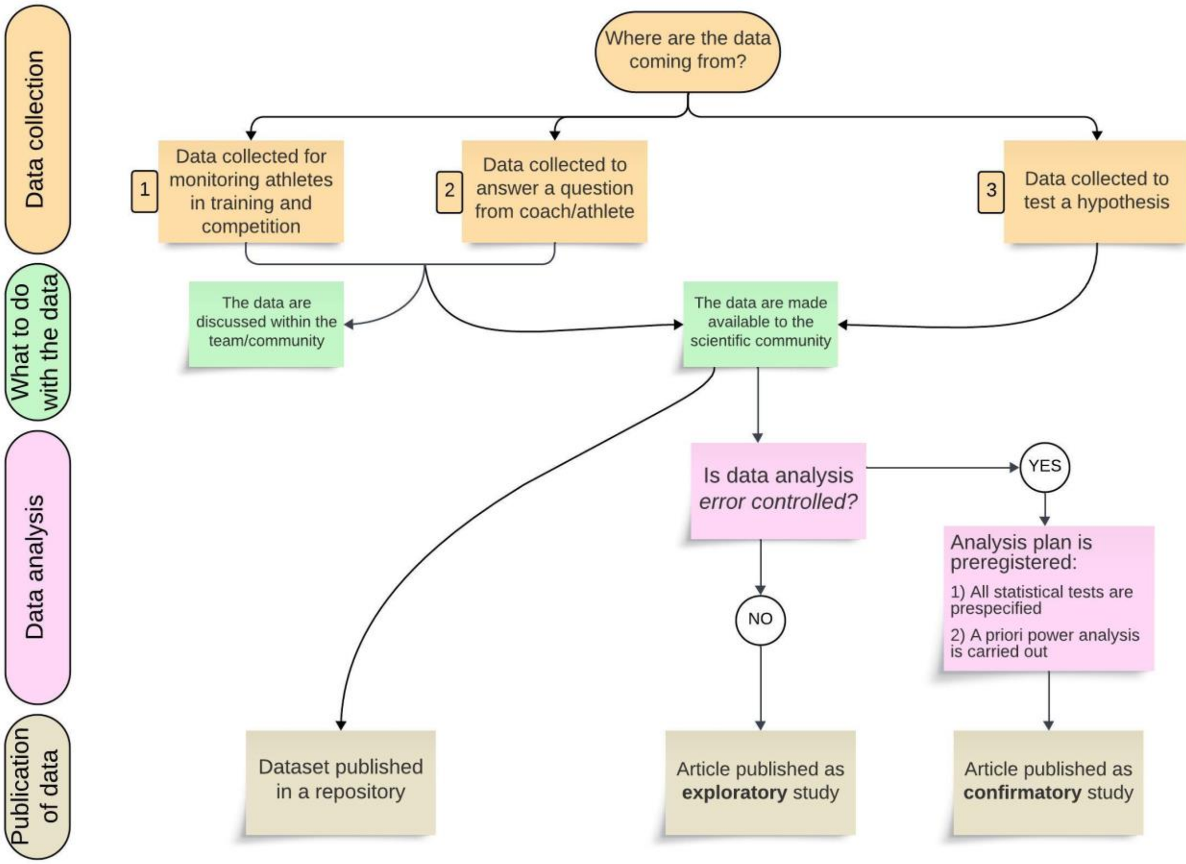

Figure 1. Flowchart showing the steps from data collection to publication, highlighting the distinct pathways of exploratory and confirmatory studies.

Data are collected for multiple reasons. First, in competitive sport, athletes' performance is typically monitored regularly during training and competition (case 1). Depending on the data to be collected and how easy/time consuming data collection is, this typically occurs on a daily (e.g. well-being scores), weekly (e.g. GPS metrics), monthly (e.g. vertical jump performance), or quarterly (e.g. sprinting abilities) basis. For instance, Olympic sport training centres can collect data from a wide range of athletes on different performance tests and then analyse the data retrospectively (e.g. Haugen et al., 2021). A dataset of performance tests may already be available because it has been collected over the years by a coach with their athletes or by a research faculty with their sport and exercise science students. It is important to highlight that in case 1 data are primarily collected to monitor athletes, inform training, and analyse matches/competitions, not to carry out a scientific study and test hypotheses. Occasionally, there may be a coach or an athlete wanting to answer a specific training question (case 2). For instance, a swim coach may wonder what happens to the level of neuromuscular fatigue if a particular type of high-intensity sprint training is administered twice instead of once a week to their mid-distance swimmers. Will they all respond in the same way? Will their ability to tolerate steady-state training improve? Will their swimming performance improve? In this case, data will be collected with a specific question in mind, but there is no explicit consideration of the study design. All data are collected on the available athletes, but there is no explicit consideration of the sample size needed to control error rates (i.e., to have high statistical power when performing tests of hypotheses). Finally, data are sometimes collected to carry out a confirmatory study (case 3) designed to test one or more prespecified hypotheses and answer research questions.

Interrogating the data

Typically, the data collected are just used within the team or within the community of coaches to make sense of training progression or difficulties that may arise (Figure 1, cases 1 and 2). However, the data can also be made available to the scientific community. The easiest way to do so is to publish the dataset following the Findable, Accessible, Interoperable, Reusable (FAIR) principles (Wilkinson et al., 2016) in an open repository such as Dryad or Figshare with a citable DOI for others to re-use. A health-related example is the UK Biobank (https://www.ukbiobank.ac.uk/), which is a database containing information from over 500,000 individuals for public health research and can be accessed through an application for a specific research project. Comparable to Biobank is the Open Science data repository by NASA to study the impact of space travel on the human body (https://osdr.nasa.gov/bio/index.html).

Similarly, it would be beneficial if more sports governing bodies and Olympic centres shared their data in open repositories for the sake of exploration. We are however aware of the ethical issues with sharing data from professional sports teams. Even if players give consent, the data often contractually belong to the team which may not be willing to make the data available to others (Ramírez-López et al., 2021; Australian Academy of Science, 2022). To break the status quo the concept of 'coopetition' has recently been introduced (Ramírez-López et al., 2021). It implies the coexistence of competition and cooperation between sports teams and national squads, as they are rivals on the field but at the same time face similar challenges, such as reducing injury risk. Coopetition can be best implemented with the involvement of a third party. A university in the role of third party would provide academic expertise that could help identify shared challenges, manage ethics approval of research projects, help design robust protocols for data collection, and disseminate results; etc (Ramírez-López et al., 2021).

Another option is that the data are examined more formally with the final goal of publishing a paper in a scientific journal. In this case a distinction can be made between the 'context of discovery' and the 'context of justification' (Reichenbach, 1938). In the context of discovery data are explored to generate new ideas, while in the context of justification, hypotheses are required to be severely tested. A test is severe when it has a high probability of showing that the hypothesis is true when it is really true, and false when it is really false. In frequentist hypothesis testing, to conduct a severe test researchers must control the Type I (α, false positive) and Type II (β, false negative; 1 – the statistical power of the test) error rates (Lakens, 2019). Thus, whenever Type I and Type II errors are higher than their prespecified level, a test will be non-severe. As an example of a non-severe test, if researchers test a hypothesis without controlling the Type II error rate, the statistical power of the test is likely to be low, and the test is likely to yield a non-significant result even if there is an effect.

In the context of discovery researchers typically do not (or cannot) perform severe tests. When analysing data researchers might take a data-driven approach to the analyses they perform. One test result might inspire a subgroup analysis in a follow-up test, which would not have been performed if the first test had showed a different result. This flexibility in how data are analysed can be useful when the goal is to inspire new hypotheses, but it inflates the Type I error rate and increases the possibility of misleading claims (or false positive results) when the goal is to severely test hypotheses. A lack of error control is one of the underlying factors contributing to the challenges researchers have in replicating published effects across many scientific fields (Murphy et al., 2023).

Error control is typically only possible when a study is designed by a scientist with the purpose of testing a specific hypothesis before collecting data (Figure 1, case 3). The Type I error rate is controlled if the researcher can specify all the tests they will perform to justify a single scientific claim. The Type II error is controlled by performing an a priori power analysis to detect the effect of interest. All the analyses must be preregistered before data collection starts (Nosek et al., 2018). For instance, a researcher might hypothesise that the consumption of a particular supplement before an all-out effort improves time trial performance. Thus, the researcher designs a study to test their hypothesis, prespecifies an alpha level and collects sufficient observations to have high statistical power to detect the effect of interest. This type of study typically lends itself to confirmatory – or error controlled – tests of a hypothesis and may be published as a confirmatory study. When researchers instead want to flexibly analyse the data to discover patterns or associations, exploratory – or non-error controlled – tests on the data are more appropriate, and the results can be published as an exploratory study.

Defining exploratory and confirmatory hypothesis testing

There are two conceptual approaches to defining confirmatory and exploratory research. On the one hand, research can be treated as confirmatory if there is a clear derivation chain from a scientific theory to a statistical hypothesis, and exploratory if a hypothesis is not derived from a theory, but instead based on for example a hunch. On the other hand, research can be treated as confirmatory if the hypothesis is severely tested from an error-statistical perspective, and exploratory if the error rates are not controlled (Lakens, 2024). In this article we take this second perspective. Thus, a fundamental trait of confirmatory studies is their ability to make claims based on hypothesis tests which have a controlled maximum probability of leading to an error, i.e. by controlling both Type I and Type II error.

Based on a statistical test, researchers can claim an effect is present, or absent. Making error-controlled claims is a unique goal of frequentist hypothesis tests, and many sport and exercise scientists seem to want to make claims in their scientific papers. In confirmatory studies, the Type I and Type II error rates should be as low as the researcher can afford to make them, and they must be prespecified before data collection. In exploratory research neither type of error can typically be controlled. First, when researchers do not report all tests they perform, the Type I error rate will be inflated. Furthermore, the choice of test to perform is often data-driven, i.e. the data guide researchers to exploratory test results that they believe yield interesting results, and consequently the probability of finding a significant result in the long run cannot be computed. For example, a researcher might decide to perform subgroup analyses only when there was no difference between conditions in the main analysis (e.g. cycling vs running training), but only for those subgroups that after visualising the data showed most variability (e.g. < 5 years training experience). As the decisions underlying the choice to perform analyses cannot be modelled, the Type I error rate cannot be controlled. This problem is further amplified by the often-vague research questions in exploratory analyses (i.e., vague questions allow for more flexible exploration).

It is similarly difficult to control Type II errors in exploratory analyses, but mainly because there is often no prespecified effect of interest that the study has been designed – and powered – to detect. When the sample size is not determined based on an a priori power analysis and the final sample size is small, the Type II error rate for effects of interest can be high. This problem is not present when exploratory tests are performed on very large datasets, as these tests would have high power for most effect sizes of interest. It is also possible that the Type II error rate in exploratory analyses decreases due to the inflation of the alpha level. As the statistical power is dependent on the alpha level (i.e., the higher the alpha level, the higher the statistical power), exploratory analyses increase the probability of finding true effects if there are true effects to be found, at the expense of an increase in false positives.

To summarise, in exploratory data analysis error rates are not controlled, (perhaps except for Type II error rates in very large datasets), while in well-designed confirmatory studies Type I and Type II error rates are controlled. Therefore, there is a higher chance that claims based on exploratory analyses are incorrect, although it is impossible to known how often they will be incorrect in the long run. This is the reason why patterns and effects identified in exploratory analyses need to be tested in a confirmatory study before researchers can claim they have observed an effect. Researchers should not make claims based on exploratory analyses, because the probability that these claims are wrong can be unacceptably high.

Additional considerations

In practice, studies can have a mix of confirmatory and exploratory tests (de Groot, 2014; Gaus et al., 2015), and some tests can control error rates more strictly than others. A study might have a primary hypothesis, which is the main goal of the study, and for which the Type I and Type II error rates are controlled. Then, the study might have a secondary hypothesis, where the Type I error rate is controlled, but for which the Type II error rate is not controlled (as the sample size is determined based on the required effect size for the primary hypothesis). For instance, a researcher may want to carry out a study where they examine the effect of a resistance training intervention on muscle fibre discharge rate of the vastus lateralis muscle in amateur soccer players. They may want to test the primary hypothesis that the alteration in discharge rate due to the training intervention is large enough to matter. In this case the researcher would specify the smallest effect size that they consider meaningful and conduct an a priori power analysis to determine the minimal sample size of the study. They also formulate a secondary hypothesis, where they explore the effect of the training intervention on muscle architecture of the same vastus lateralis. In such case, the researcher still controls the Type I error at the prespecified level but not the Type II error. Additionally, exploratory tests on other variables, such as recruitment and derecruitment thresholds, are reported. Here, researchers explore subgroup analyses, covariates, and a range of dependent variables. They let their choice of analyses depend on previous test results. For these analyses, neither the Type I nor the Type II error rates are controlled.

Besides controlling for Type I and Type II error rates, a purely confirmatory test is only as informative as the quality of the study. If a test has low internal and/or external validity, it is irrelevant even if the error rate is controlled. This is especially relevant in situations where researchers need to deviate from a preregistered confirmatory test (Lakens, 2024). If a deviation from a preregistered analysis plan is necessary, the error control is slightly reduced, although typically not as much as for exploratory tests.

Finally, the reader may wonder where descriptive and observational research, as well as pilot/feasibility studies, fall within the exploration-confirmation continuum. In descriptive research no tests are performed, no claims about the presence or absence of effects are made, there is no error control, and instead of inferential statistics, descriptive statistics are reported. Observational studies are quite common in sport and exercise science. There might not be a manipulation of independent variables, but inferential statistics are used to test how one or more dependent variables differ between categories or conditions. While discussing the Applied Research Model for the Sport Sciences, Bishop (2008) advocates that descriptive/observational studies can provide valuable insights into the other stages of the model, which includes controlled laboratory and field studies. Although not necessarily so, observational studies might in practice fall closer to the exploratory end within the exploration-confirmation spectrum. Pilot/feasibility studies are designed to test out the equipment, refine the methods of data collection so that any problem can be spotted early enough and possibly ironed out, and to assess the ability to recruit/retain participants (Scheel et al., 2021). The often-small sample size means hypothesis tests will have Type II error rates that are too high to be informative, and descriptive statistics such as means and effect sizes will have large uncertainty (i.e., wide confidence intervals around the estimate of interest).

WHAT MAKES AN EXPLORATORY STUDY VALUABLE AND INTERESTING?

One reason researchers present exploratory tests as confirmatory tests is because the latter are perceived to be more valuable and interesting. Confirmatory tests can be very valuable, as they allow researchers to make error-controlled claims. But such studies are often only possible if they build on a solid foundation provided by high quality exploratory work. In this section we aim to provide some suggestions to clarify the value of a transparently reported exploratory study. Of course, many of these factors also make confirmatory research more valuable, but they are especially useful to provide an argument for why exploratory research is worth publishing.

1) *Large sample sizes to minimise random variation*. If the goal of exploration is to detect if there is anything interesting in the data, then the larger the sample size, the better, because the Type II error rate will be lower. The question when the sample size is large enough depends on the smallest effect the researcher is interested in. Beyond exploratory testing, large datasets have additional benefits, such as the opportunity to derive accurate estimates that could provide input for future hypothesis tests. For instance, for a particular measure of interest it may be possible to accurately estimate the standard deviation of the measure, which can subsequently be used for an a priori power analysis for the confirmatory test. Yet, the sample size for exploratory studies is typically not based on an a priori power analysis. It is therefore useful to conduct a sensitivity power analysis to provide insights into the effect sizes that the study design has sufficient power to detect when performing exploratory tests (Lakens, 2022).
2) *Uncommon sample composition*. Data coming from a population that is not easily accessible is valuable, as it greatly increases our knowledge about the generalizability of existing knowledge from these populations. For example, top-level athletes; athletes from a less popular/extreme sports; athletes with a particular type of injury history; etc. Such datasets can provide a rough indication about whether effects observed in other populations generalise or provide a first indication of the possibility that there are boundary conditions.
3) *Rigorous data collection*. Whoever collects the data should be trained to do so. Data collection should follow validated and detailed protocols. Measurement error should be minimised. While some level of random error is inevitable, it is critical to be vigilant about systematic error (Tong, 2019). Given that in exploratory research the data were originally not collected for the purpose of scientific research, this factor is even more important. Additionally, if the person using the data did not collect them, they must exercise care as the quality of the data may be unknown.
4) *Widely used measures, and the identification of patterns, associations and clusters that are meaningful for the scientific community*. The more standardised and established the measurements are, the better, because the data are then also valuable for other researchers. For example, in tennis, finding an association between number of unforced errors and fatigue due to playing time will inform training practice; whereas a one-off association between handedness and number of unforced errors may be eye-catching but most likely a fluke of little practical value.
5) *Observing a logical and coherent pattern across multiple variables*. For example, if multiple related physiological variables (e.g. heart rate, VO2, respiratory rate, blood pressure, etc) measured on the same participants follow a coherent pattern over time, it reduces the probability of exploratory claims being a fluke. If only certain unrelated variables that theoretically should

reveal similar effects all point in the same direction, the confidence in the claims is increased even more. For instance, if maximal isometric voluntary contraction of quadriceps improved the general ability of sprinting and jumping during a soccer match, while reducing the injury risk, then it may be associated with acceleration, deceleration, change of direction, distance covered at high speed, number of successful headers, and playing time. This can be further emphasised by using triangulation (Munafò & Smith, 2018; Tong, 2019), whereby the data are analysed using different methodologies. If the results substantially agree, they are less likely to be caused by random variation, and a true effect becomes more likely.

6) *The potential for research question generation*. By freely observing and exploring the data, incidental new findings may be revealed that can lead to one or more new research questions and relative hypotheses. These hypotheses will then have to be severely tested by new data in a confirmatory study. For example, in a large dataset from the shot-put event a researcher may observe that there is a positive relationship between the distance thrown and the eccentric force developed during the downward phase of a countermovement jump. The researcher may use this pattern to develop the following research questions: (a) Would a training-related improvement in eccentric force drive an increase in shot put performance? (b) How would that compare to a training programme focused on improving the force expressed during the stretch-shortening cycle? In other words, is it really the eccentric phase the coach needs to focus on, or could similar/better results be achieved by targeting the stretch-shortening cycle, which includes both eccentric and concentric action? Would that work to the same extent and in similar fashion for males and females?

It is also worth pointing out when exploratory analyses are less interesting. Imagine a researcher who has access to a dataset that contains data from male and female elite and highly trained athletes, with a wide age range. Some of them may have trained at altitude, which can be interesting to compare against those who did not complete altitude training. As an example of a good exploratory analysis, the researcher could test for subgroup differences and correlations, running as many tests as they wanted. While by definition exploration is free and mostly depends on the aim of the analyses, some choices for exploratory analyses are less justifiable. For instance, although it may be interesting to explore what causes outliers in the dataset, there are no compelling arguments that justify the selective removal of participants on a correlation between variables, especially if that decision is only based on whether the correlation analysis is statistically significant. Similarly, it is rarely an interesting question to explore what will happen if a few additional datapoints are collected and added to the relationships between variables. In confirmatory analyses, such practices, especially if undisclosed, are seen as ways to change non-significant into significant test results, which is an improper research practice. Although in exploratory research there are no strong objections to such data analysis strategies, and it might be interesting to search for relationships with large effects or low p-values, when this becomes the main goal of exploratory analysis, its value and interest is greatly reduced. Importantly, the researcher should be aware that exploratory tests do not have sufficient error control, and therefore they will be unable to make error-controlled claims. Test results can only be used as a tool to help formulate hypotheses for future confirmatory studies.

CARRYING OUT AND WRITING AN EXPLORATORY STUDY

Perhaps the most important change that is needed in the sport and exercise science discipline is that exploratory tests are transparently reported. The researcher should acknowledge where the data are coming from, and that data-driven analyses have been carried out with no or insufficient error control. We propose that the exploratory nature of the study should at the minimum be reported in the abstract. The authors may choose to have 'exploratory' in the title of the paper as well. This is in line with recommendations by Janiszewski and van Osselaer (2021) for exploratory research in consumer psychology. Similarly to the PRISMA guidelines for systematic reviews (Page et al., 2021), including this information will facilitate the identification of the type of study by potential users. The language should reflect the nature of the study throughout the paper (see below section 'Writing up and submitting for publication').

Data analysis

In previous sections we have clarified the difference between confirmatory and exploratory hypothesis testing. This will be further discussed here as a key method of data analysis in exploratory research (Gaus et al., 2015). However, exploratory data analysis also encompasses methods to summarise and visualise the data, e.g. to identify patterns, associations between variables, interactions, clusters of individuals, or trajectories over time (Höfler et al., 2023).

Hypothesis testing. In exploratory hypothesis testing researchers perform statistical tests that have a higher error rate than confirmatory tests. Due to the data-driven selection of the tests while searching for interesting results, significant results will be observed at a (possibly much) higher rate than the alpha level, even if there are no true effects in the population.

Exploratory hypothesis tests return the same quantitative results – p-values and effect sizes – as confirmatory tests. The only difference is how exploratory hypothesis tests are interpreted and presented. It is crucial to realize that researchers should not make claims based on exploratory tests (Gaus et al., 2015). In a confirmatory test researchers can make error-controlled claims, e.g. “we observed a statistically significant effect of caffeine on reaction time during sprint performance, which we will interpret as the presence of a true effect, while acknowledging that in the long run, at most 5% of such claims are a false positive”. The idea is that a 5% probability of an error is deemed low enough to act as if there is a true effect – at least until future data suggests otherwise. Statistical tests can be seen as a methodological procedure to turn observed data into scientific statements with an acceptably low error rate (Tunç et al., 2023). Exploratory tests lack the severity of the methodological procedure that underlies confirmatory tests (Mayo, 2018).

It might be tempting to make a claim if the observed p-value is very low (e.g., $p = 0.0002$), as the test would have been able to withstand strict corrections for multiple comparisons. A lower p-value will be correlated with a higher probability that there is a true effect when there is a mix of true and null effects. At the same time, it is important to appreciate that when there is no effect, statistical tests will return p-values that are uniformly distributed, which means that each p-value (e.g., $p = 0.0002$ and $p = 0.9998$) is equally likely to be observed (Mesquida et al., 2022). It is important to realise that claims based on exploratory tests might well be true (and in general, the lower the p-value, the higher that probability will be in the long run), but such claims are not warranted based on a coherent methodological procedure to translate noisy empirical observations into reliable scientific claims.

If claims that are not error controlled are not warranted, then what is the added value of exploratory hypothesis testing? The main function of exploratory tests is a tool to update subjective beliefs about which effects are worth exploring in the future. The same information from a study that is used to compute a p-value or an effect size can be used to compute a Bayes factor, given a specific prior (Francis, 2017). A Bayes factor from an exploratory test cannot be used to make an error-controlled claim, but it can be used to inform peers about how much they should update their prior belief in a hypothesis or how much a researcher should update their belief in one hypothesis relative to another hypothesis (e.g., the null hypothesis relative to a specific alternative hypothesis). For instance, by analysing 2229 bouts in mixed martial arts, it was recently reported that the anecdotal assumption of stature and arm span being an advantage for athletes returned a Bayes factor for the null hypothesis of 11 and 7, respectively. Therefore the prior belief was updated as athlete morphology appears to be irrelevant for mixed martial arts success (Kirk, 2024). The higher the posterior belief in a hypothesis, the more interesting it is to study in a subsequent confirmatory experiment. In other words, exploratory analyses can be used to update the subjective belief of researchers in how interesting it is to examine a hypothesis in the future.

In exploratory studies researchers need to report the reasons for why they performed each exploratory analysis. An excellent way to achieve this is through an open lab notebook (Figueiredo et al., 2022). At the start the researcher writes down a complete log of every analysis, starting with the motivation for why it was performed, the result of the exploratory test, and the inferences they drew. Each analysis may be inspired by a strong prior based on the literature, results from previous exploratory analyses, a hunch, or can simply be performed because it is possible, and all other plausible tests have already been performed. Although different researchers will have different subjective priors for each hypothesis, a transparently documented log of why exploratory analyses were performed will provide useful information to base priors on, and the more random an exploratory test is, the lower the prior (and hence, the lower the posterior belief). Such transparently reported exploratory analyses can on average prove useful in deciding which effects to examine in follow-up error-controlled studies.

Data summary and visualisation. Summarising the main characteristics and visualising the data is also part of the exploration process. Descriptive statistics are used to identify central tendency (e.g. mean and median) and dispersion (e.g. variance, standard deviation, quartiles, range) of a dataset, but also extreme values and potential outliers (e.g. by examining min and max values in a boxplot or scatterplot).

Clarity of data presentation and visualisation with the use of tables, graphs and maps is essential in exploratory research and it is nowadays a fundamental skill (Cortex, 2021). For example, the pattern of VO2 data during a steady state exercise, or the trajectory of well-being scores over a 6-month period can be presented using a line graph; a scatterplot can show the association between muscle force and muscle thickness. The strength of the association may or may not vary across the range of muscle force values; clusters of individuals playing the same sport and showing similar daily calorie intake may become visible using clustering graphs; boxplots can be used to summarise central tendency and dispersion of knee-extension and knee-flexion torque of athletes attending an Olympic training centre; line graphs would display the interaction between resistance training experience and number of training sessions per week when it comes to examining the improvement in 1RM; etc.

The examination of patterns becomes more complex when trying to fit the data into a model. For example, it is well known that the kinetics of VO2 fits an exponential curve when the exercise intensity rises from 0 to a submaximal steady state level. When the researcher wants to explore

the VO2 kinetics of several athletes, it may be useful to examine the pattern of the residuals (Tong, 2019), i.e. how much each observed datapoint deviates from the theoretical value within the exponential relationship. It may be that residuals are smaller or larger in the first section of the curve for athletes with greater/lower VO2max, etc. Similarly, the vertical ground reaction forces measured during the stance phase of gait fit a typical waveform with two peaks and one trough in the middle. By examining the pattern of residuals, excessive deviations may be a sign of a particular type of shoe, of an injury, of a pathology, etc.
Data reduction techniques are common in exploratory research. Even though there may be similarities between some of them, we have briefly described below those that are typically used in our field. In a complex dataset *principal component analysis* allows the removal of redundant variables that are highly correlated and dependent to one another and therefore do not provide additional information (Weaving et al., 2019). One example of application of this technique in sport and exercise science is the identification of key performance indicators in swimming (Staunton et al., 2024).
Factor analysis can be used to identify a set of variables with a similar pattern of variation (i.e. covariation) and relate them to one common latent construct or factor (Bandalos & Finney, 2018). For instance, talent in a particular sport could be the factor that derives from the covariation of 5 variables, e.g. genetic marker a, genetic marker b, height, body fat percentage, and lower limb muscle strength. The purpose of exploratory factor analysis is to detect pattern similarities, extract factors and generate hypotheses about their structure (Costello & Osborne, 2005; Watkins, 2018). In our field factor analysis has been used to identify the factors that explain fitness status in youth football players within variables of physical performance, anthropometry and biological markers (Perroni et al., 2023), just to provide one example.
Cluster analysis is another technique of data reduction that aims to identify similar observations, or individuals with similar features, and put them into clusters. As an example, it has been used to characterise sleep behaviour in elite athletes (Suppiah et al., 2022), or to categorise kinematic patterns in runners based on injury status and location (Jauhiainen et al., 2020). Free software packages like Jamovi and JASP support principal component analysis, exploratory factor analysis and cluster analysis, so does R and other commercial statistical packages like SPSS.
For more insight on data exploration the reader is referred to Pearson (2011). For those using R, or wanting to learn R, there is an interesting course entitled “Exploratory Data Analysis: The basics of exploring data in R”, which is built into the package ‘Swirl’ and encompasses 14 lessons (Kross et al., 2020). There is also an online textbook worth reading (Nordmann & DeBruine, 2023).

Writing up and submitting for publication

In Table 1 we outline a suggested structure for an exploratory paper.

Table 1. Recommendations for reporting an exploratory study

Section	Description
Title	The title may include the word 'exploratory', such as "An exploratory study of…".
Abstract	The abstract **must** report the exploratory nature of the study.
Introduction	Introductions can be short, or even included in the Methods section if the existence of a dataset was the starting point for the study. Instead of an introduction, a section ***Background and Methods*** could provide background information on the dataset and the context of data collection. The introduction should explain what makes the exploration valuable, can consider the size and composition of the sample of participants, and why conducting such exploratory study is of interest for the community. Importantly, the Introduction does not follow the typical structure of a confirmatory study as there is no mention of hypotheses to be tested.
Methods	It should include ethical considerations and detailed information on participants, equipment/materials, measurement protocols, data processing, variable extraction, data/statistical analysis. In these regards it does not differ from a confirmatory study. As exploratory studies are not performed with a specific hypothesis test in mind, the sample size justification cannot be based on an a-priori power analysis. However, it is important to report a sensitivity power analysis or the minimum detectable effect size, to indicate which effect sizes could be detected in exploratory tests. The dataset and code for the analyses should be shared publicly (e.g., Dryad or Figshare) or included in appendix following the FAIR principles.
Results	Results of **all** exploratory analyses **must** be reported. Reporting of findings should be stated at the data level, and not as claims that generalize beyond the data. E.g. "There is a correlation between a and b, and there is a difference between c and d", but not 'Athletes perform better under A than B'. Non-significant results do not mean there is no effect at all – the effect might just be smaller than could be detected given the sample size of the study.
Discussion	Discussion is limited and may well be included in the Results section, which could be labelled as ***Results and Discussion***. Observing patterns, correlations, differences, etc. should not be presented as a scientific claim of the presence or absence of an effect, but as observed patterns in the data that will be used to formulate hypotheses to be severely tested in future studies.
Conclusion	It **must** include a sentence along these lines: "All potential effects identified in this paper need to be confirmed in a future study with adequate error control and statistical power". It should also suggest which hypotheses would be worth investigating in a future confirmatory study.

FAIR = Findable, Accessible, Interoperable, Reusable (Wilkinson et al., 2016)

As discussed above, at least the abstract and possibly the title of the paper should clarify the exploratory nature of the study. Transparency from the outset about the exploratory nature of the paper would benefit the researchers, the reviewers, and the readers (McIntosh, 2017). Although perhaps contentious, and it might be too early to implement in practice, it is possible that we are moving to a future where, without obvious signs that hypothesis tests are confirmatory (and preregistered), studies are treated as exploratory.

In partial agreement with Höfler et al. (2023), we believe that Introduction and Discussion of an exploratory paper do not have to follow the same structure as a confirmatory paper or may be missing altogether. While Methods and Results will be by far the largest sections, most likely some background information to explain the context of data collection is relevant and should be included. Similarly, basic discussion around what the results show may well be done within the Results section. This is also in line with guidelines for writing an exploratory report for the journal 'International Review of Social Psychology' (https://rips-irsp.com/exploratory-reports#exploratory-reports-at-irsp-guidelines-for-authors). Interestingly, in relation to reporting exploratory research, Janiszewski and van Osselaer (2021) advocate the introduction of a 'flexible reporting template', which should be acknowledged by review teams, to increase the value of data exploration and lessen the importance of existing theory.

Results of all exploratory analyses should be reported, and researchers should not selectively report only strong associations or patterns, or significant results (Luijken et al., 2022). Reporting of findings should be stated as they are, without any emphasis, e.g. "there is a correlation between a and b, and there is a difference between c and d". Importantly, at the end of the Results or Discussion (if present) section, there should be a concluding statement that reads approximately: "All potential effects identified in this paper need to be confirmed in a future study with adequate error control and statistical power". To reiterate, observing a pattern or an association in the data should not be reported as a new finding, but it should encourage the researcher to set up a new pre-registered study with a clear plan, so that what was observed during exploration can be severely tested in a follow-up confirmatory study (McPhetres, 2020). The conclusion of an exploratory study should propose a research plan that reflects the exploratory results. Rather than a generic "more research is needed in this area", it should suggest what would be worth investigating and how in a future confirmatory study (Luijken et al., 2022). For instance, "of all exploratory analyses performed, it is striking how female athletes did not seem to change their pattern of eccentric force during a vertical jump because of sprint training, which is well known to change in males. Therefore, an adequately powered confirmatory study should examine the effect of 8 weeks of sprint training on the pattern of eccentric force during a vertical jump in female athletes".

Finally, it is important to ensure that the dataset is shared publicly whenever possible following the FAIR principles (Wilkinson et al., 2016) as mentioned above in 'Interrogating the data'. Both raw data and the code for the analyses conducted should be published with the paper to allow other researchers to explore the dataset (Höfler et al., 2023).

When it comes to submitting an exploratory study for publication, choice of journal and guidelines are not straightforward. A few years ago, the journal 'Cortex' created a new article type – exploratory reports (Cortex, 2021; McIntosh, 2017). The same happened in 2018 for the journal 'International Review of Social Psychology' (IRSP, 2018). Unfortunately, these initiatives failed as the number of submissions has been extremely low for either journal. The only journal in sport and exercise science explicitly accepting exploratory research is ‘Science and Medicine in Football’. In the category Original investigations, they state: "Exploratory studies are accepted, but their nature should be transparently reported in the manuscript. While they can be used to generate hypotheses, they are not intended for making recommendations or practical applications". With the assistance of a shiny app (DeBruine & Lakens, 2024) we examined all the papers accepted in this journal (n=440) that included the words ‘explor*’ or ‘exploratory research’ somewhere in the text (search conducted on the 23rd May 2024). We found that four papers used some exploratory techniques of data reduction, such as factor analysis and principal component analysis; only one study stated that the analyses were exploratory, with the aim of identifying hypotheses to be tested in future studies. One study did not formulate a hypothesis because of its exploratory nature; another study tested two hypotheses but also explored other variables; two studies did not adjust the alpha level for multiple statistical tests because of its exploratory nature and the low cost in incurring an error. We are not aware of any other Sport and Exercise Science journals specifically accepting exploratory studies, but that does not mean that they cannot be submitted and accepted for publication if they are well conceived, performed, and written.

‘Scientific Data’, published by Nature, is a journal for “descriptions of datasets and research that advances the sharing and reuse of research data” (https://www.nature.com/sdata/). Even though it doesn’t explicitly mention exploratory research, there are similarities with what we have described, including the format of the article as we discussed above in ‘Writing up’. For a sport and exercise science example published in this journal the reader is referred to Derlatka and Parfieniuk (2023).

Due to their data-driven nature of asking research questions, in line with de Groot (2014) we think that exploratory studies do not need to be pre-registered (Lakens et al., 2024). There are however other authors who claim that at least some level of preregistration is needed for exploratory research (Dirnagl, 2020; McPhetres, 2020). Given that preregistration of confirmatory studies is still far from becoming widespread practice in our field, and the logical arguments against preregistering exploratory studies, we believe we should focus on preregistering confirmatory research.

In an ideal world it would be highly desirable to link confirmatory follow-up studies to the exploratory research that inspired the confirmatory test in scientific publications. It would tell a story of how a hypothesis has been exploratively generated and then subsequently tested in a confirmatory study. Interestingly, for clinical research with animal studies, Mogil and Macleod

(2017) have proposed a new type of paper combining a confirmatory study with the exploratory work that preceded the final study.

CONCLUSIONS AND FINAL REMARKS

In this paper we have made the case for exploratory research and why it is relevant and important in the sport and exercise sciences. From experience, many investigations in our field use 'intransparent exploration' which leads to a higher rate of false claims. With the help of a flow diagram (Figure 1), we have described the features of exploratory and confirmatory research, with the ability of controlling error rate being the dividing line between the two types of research, and we have provided guidelines to decide on whether a study is about exploring the data, or confirming a hypothesis, or a bit of both. We have emphasised that there is nothing wrong with an exploratory study if, to quote Marcia McNutt, former editor of Science (Shell, 2016), "scientists call it as such and do not try to pass it off as something else". As Schwab and Held (2020) state: "Finding the questions to ask is at least as crucial as answering them, if not more so". We have also highlighted what makes a good exploratory study, how to carry out exploratory data analysis, and how to write up exploratory results.

We would like to conclude with Tong (2019), who states that most scientific research is exploratory in nature and should be examined using descriptive and exploratory analysis, which includes summary statistics, statistical graphics and tables, and disciplined data exploration. However, most of the statistical teaching is overfocused on statistical inference (Tong, 2019). The latter statement is unfortunately true also in our field. Sport and Exercise Science degrees should incorporate early in their curriculum a module including data handling, data description and exploratory data analysis. It should teach students good 'statistical thinking', i.e. how the data relate and what they can tell about real-world problems; what is the context and what are the constraints (Tong, 2019). We need to educate scientists about the value of properly conducted and transparently reported exploratory research. One of the best ways to achieve this is to give exploratory research the place in our scientific literature that it deserves.

Contributions

Contributed to conception and design: MD, CM, DL
Drafted and/or revised the article: MD, CM, GA, DL
Approved the submitted version for publication: MD, CM, GA, DL

Acknowledgements

We would like to acknowledge the insightful discussion on exploratory research we had with Sajedeh Rasti, Dr Paul Riesthuis and Laura Sita.

Funding information

N.A.

Data and Supplementary Material Accessibility

N.A.

REFERENCES

Australian Academy of Science. (2022). Getting Ahead of the Game: Athlete Data in Professional Sport. www.science.org.au/datainsport

Bandalos, D. L., & Finney, S. J. (2018). Factor Analysis. Exploratory and Confirmatory. In G. R. Hancock, L. M. Stapleton, & R. O. Mueller (Eds.), The Reviewer's Guide to Quantitative Methods in the Social Sciences (pp. 98-122). Routdlege.

Bishop, D. (2008). An applied research model for the sport sciences. Sports Medicine, 38(3), 253-263. https://doi.org/10.2165/00007256-200838030-00005

Cortex. (2021). Exploratory reports - author guidelines. Retrieved 12 April from https://legacyfileshare.elsevier.com/promis_misc/Exploratory_Reports_Guidelines.pdf

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. Practical assessment, research, and evaluation, 10(7).

de Groot, A. D. (2014). The meaning of "significance" for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han L. J. van der Maas]. Acta Psychologica, 148, 188-194. https://doi.org/10.1016/j.actpsy.2014.02.001

DeBruine, L., & Lakens, D. (2024). papercheck: Check Scientific Papers for Best Practices. R package version 0.0.0.9014. Retrieved 28 May from https://scienceverse.github.io/papercheck/, https://github.com/debruine/papercheck.

Derlatka, M., & Parfieniuk, M. (2023). Real-world measurements of ground reaction forces of normal gait of young adults wearing various footwear. Scientific Data, 10(1), 60. https://doi.org/10.1038/s41597-023-01964-z

Dirnagl, U. (2020). Preregistration of exploratory research: Learning from the golden age of discovery. Plos Biology, 18(3), e3000690. https://doi.org/10.1371/journal.pbio.3000690

Figueiredo, L., Scherer, C., & Cabral, J. S. (2022). A simple kit to use computational notebooks for more openness, reproducibility, and productivity in research. PLoS Computational Biology, 18(9), e1010356. https://doi.org/10.1371/journal.pcbi.1010356

Francis, G. (2017). Equivalent statistics and data interpretation. Behavior Research Methods, 49(4), 1524-1538. https://doi.org/10.3758/s13428-016-0812-3

Gaus, W., Mayer, B., & Muche, R. (2015). Interpretation of statistical significance-exploratory versus confirmative testing in clinical trials, epidemiological studies, meta-analyses and toxicological screening (using Ginkgo biloba as an example). Clininical & Experimental Pharmacology, 5(4), 1000182.

Haugen, T. A., Breitschädel, F., Wiig, H., & Seiler, S. (2021). Countermovement Jump Height in National-Team Athletes of Various Sports: A Framework for Practitioners and Scientists. International Journal of Sports Physiology and Performance, 16(2), 184-189. https://doi.org/10.1123/ijspp.2019-0964

Höfler, M., McDonald, B., Kanske, P., & Miller, R. (2023). Means to valuable exploration II: How to explore data to modify existing claims and create new ones. Meta-Psychology, 7, *MP.2022.3270*

Höfler, M., Scherbaum, S., Kanske, P., McDonald, B., & Miller, R. (2022). Means to valuable exploration: I. The blending of confirmation and exploration and how to resolve it. Meta-Psychology, 6, *MP.2021.2837*

Hussain, S. J., & Cohen, L. G. (2017). Exploratory studies: a crucial step towards better hypothesis-driven confirmatory research in brain stimulation. Journal of Physiology, 595(4), 1013-1014. https://doi.org/10.1113/JP273582

ICH. (1998). Statistical principles for clinical trials (E9). International Council for harmonisation of technical requirements for pharmaceuticals for human use. Retrieved 15 April from https://database.ich.org/sites/default/files/E9_Guideline.pdf

ICH. (2022). General considerations for clinical studies E8(R1). International Council for harmonisation of technical requirements for pharmaceuticals for human use. Retrieved 15 April from https://database.ich.org/sites/default/files/ICH_E8-R1_Guideline_Step4_2021_1006.pdf

IRSP. (2018). Exploratory Reports at IRSP: Guidelines for Authors. Retrieved 06 May from https://rips-irsp.com/exploratory-reports#exploratory-reports-at-irsp-guidelines-for-authors

Janiszewski, C., & van Osselaer, S. M. J. (2021). The benefits of candidly reporting consumer research. Journal of Consumer Psychology, 31(4), 633-646.

Jauhiainen, S., Pohl, A. J., Ayramo, S., Kauppi, J. P., & Ferber, R. (2020). A hierarchical cluster analysis to determine whether injured runners exhibit similar kinematic gait patterns. Scandinavian Journal of Medicine & Science in Sports, 30(4), 732-740. https://doi.org/10.1111/sms.13624

Kimmelman, J., Mogil, J. S., & Dirnagl, U. (2014). Distinguishing between Exploratory and Confirmatory Preclinical Research Will Improve Translation. PLoS Biology 12(5): e1001863. https://doi.org/10.1371/journal.pbio.1001863

Kirk, C. (2024). A 5-Year Analysis of Age, Stature and Armspan in Mixed Martial Arts. Reserch Quarterly for Exercise and Sport, 95(2), 450-457. https://doi.org/10.1080/02701367.2023.2252473

Kross, S., Carchedi, N., Bauer, B., & Grdina, G. (2020). _swirl: Learn R, in R_. R package version 2.4.5. Retrieved 26 April 2024 from https://CRAN.R-project.org/package=swirl

Lakens, D. (2019). The value of preregistration for psychological science: A conceptual analysis. Japanese Phsychological Review, 62(3), 221-230.

Lakens, D. (2022). Sample size justification. Collabra: Psychology, 8(1), 33267.

Lakens, D. (2024). When and How to Deviate from a Preregistration. Collabra: Psychology, 10(1).

Lakens, D., Mesquida, C., Rasti, S., & Ditroilo, M. (2024). The benefits of preregistration and Registered Reports. Evidence-Based Toxicology, 2(1), 2376046. https://doi.org/10.1080/2833373X.2024.2376046

Luijken, K., Dekkers, O. M., Rosendaal, F. R., & Groenwold, R. H. H. (2022). Exploratory analyses in aetiologic research and considerations for assessment of credibility: mini-review of literature. British Medical Journal, 377, e070111. https://doi.org/10.1136/bmj-2021-070113

Mayo, D. G. (2018). How Not to Corrupt Power. In D. G. Mayo (Ed.), Statistical Inference as Severe Testing (pp. 353-370). Cambridge University Press.

McIntosh, R. D. (2017). Exploratory reports: A new article type for Cortex. Cortex, 96, A1-A4. https://doi.org/10.1016/j.cortex.2017.07.014

McPhetres, J. (2020). What should a preregistration contain? Retrieved 29 April from https://osf.io/preprints/psyarxiv/cj5mh

Mesquida, C., Murphy, J., Lakens, D., & Warne, J. (2022). Replication concerns in sports and exercise science: a narrative review of selected methodological issues in the field. Royal Society Open Science, 9(12), 220946. https://doi.org/10.1098/rsos.220946

Mogil, J. S., & Macleod, M. R. (2017). No publication without confirmation. Nature, 542(7642), 409-411. https://doi.org/DOI 10.1038/542409a

Munafò, M. R., & Smith, G. D. (2018). Repeating experiments is not enough. Nature, 553(7689), 399-401.

Murphy, J., Mesquida, C., Caldwell, A. R., Earp, B. D., & Warne, J. P. (2023). Proposal of a Selection Protocol for Replication of Studies in Sports and Exercise Science. Sports Medicine, 53(1), 281-291. https://doi.org/10.1007/s40279-022-01749-1

Nilsen, E. B., Bowler, D. E., & Linnell, J. D. C. (2020). Exploratory and confirmatory research in the open science era. Journal of Applied Ecology, 57(4), 842-847. https://doi.org/10.1111/1365-2664.13571

Nordmann, E., & DeBruine, L. (2023). Applied Data Skills. v3.0. Retrieved 19 June from https://psyteachr.github.io/ads-v3/

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. Proceedings of the National Academy of Sciences of the United States of America, 115(11), 2600-2606. https://doi.org/10.1073/pnas.1708274114

Olken, B. A. (2015). Promises and perils of pre-analysis plans. Journal of Economic Perspectives, 29(3), 61-80.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S.,...Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. British Medical Journal, 372, n71. https://doi.org/10.1136/bmj.n71

Pearson, R. (2011). Exploring Data in Engineering, the Sciences, and Medicine. Oxford University Press.

Perroni, F., Castagna, C., Amatori, S., Gobbi, E., Vetrano, M., Visco, V., Guidetti, L., Baldari, C., Rocchi, M. B. L., & Sisti, D. (2023). Use of Exploratory Factor Analysis to Assess the Fitness Performance of Youth Football Players. Journal of Strength and Conditioning Research, 37(7), E430-E437. https://doi.org/10.1519/Jsc.0000000000004414

Pham, M. T., & Oh, T. T. (2021). Preregistration is neither sufficient nor necessary for good science. Journal of Consumer Psychology, 31(1), 163-176.

Ramírez-López, C., Till, K., Boyd, A., Bennet, M., Piscione, J., Bradley, S., Giuliano, P., Leduc, C., & Jones, B. (2021). Coopetition: cooperation among competitors to enhance applied research and drive innovation in elite sport. British Journal of Sports Medicine, 55(10), 519–530.

Reichenbach, H. (1938). The three tasks of epistemology. In Experience and prediction: An analysis of the foundations and the structure of knowledge (pp. 3-16). University of Chicago Press.

Scheel, A. M., Tiokhin, L., Isager, P. M., & Lakens, D. (2021). Why Hypothesis Testers Should Spend Less Time Testing Hypotheses. Perspectives on Psychological Science, 16(4), 744-755, 1745691620966795. https://doi.org/10.1177/1745691620966795

Schwab, S., & Held, L. (2020). Different worlds. Confirmatory versus exploratory research. Significance, 17(2), 8-9.

Shell, E. R. (2016). Hurdling obstacles: Meet Marcia McNutt, Scientist, Administrator, Editor, and Now National Academy of Sciences President. Science, 353, 116-119.

Staunton, C. A., Romann, M., Björklund, G., & Born, D. P. (2024). Diving into a pool of data: Using principal component analysis to optimize performance prediction in women's short-course swimming. Journal of Sports Sciences, 42(6), 519-526. https://doi.org/10.1080/02640414.2024.2346670

Suppiah, H. T., Gastin, P. B., & Driller, M. W. (2022). A Strategy to Inform Athlete Sleep Support From Questionnaire Data and Its Application in an Elite Athlete Cohort. International

Journal of Sports Physiology and Performance, 17(10), 1532-1536. https://doi.org/10.1123/ijspp.2021-0561

Tong, C. (2019). Statistical Inference Enables Bad Science; Statistical Thinking Enables Good Science. American Statistician, 73, 246-261. https://doi.org/10.1080/00031305.2018.1518264

Tukey, J. W. (1980). We Need Both Exploratory and Confirmatory. American Statistician, 34(1), 23-25. https://doi.org/10.2307/2682991

Tunç, D. U., Tunç, M. N., & Lakens, D. (2023). The epistemic and pragmatic function of dichotomous claims based on statistical hypothesis tests. Theory & Psychology, 33(3), 403-423. https://doi.org/10.1177/09593543231160112

Wagenmakers, E. J., Wetzels, R., Borsboom, D., van der Maas, H. L., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. Perspectives on Psychological Science, 7(6), 632-638. https://doi.org/10.1177/1745691612463078

Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. Journal of Black Psychology, 44(3), 219-246. https://doi.org/10.1177/0095798418771807

Weaving, D., Beggs, C., Dalton-Barron, N., Jones, B., & Abt, G. (2019). Visualizing the Complexity of the Athlete-Monitoring Cycle Through Principal-Component Analysis. International Journal of Sports Physiology and Performance, 14(9), 1304-1310. https://doi.org/10.1123/ijspp.2019-0045

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R.,...Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, 160018. https://doi.org/10.1038/sdata.2016.18