1 Incentives for Registered Reports from a risk-sensitivity perspective

<sub>2</sub> Incentives for Registered Reports from a risk-sensitivity perspective

<sub>3</sub> Registered Reports are an article format designed to reduce publication bias and
<sub>4</sub> 'questionable research practices' (QRPs), which distort the published record of research
<sub>5</sub> findings in many scientific disciplines (Chambers, 2013; de Vries et al., 2018; Dickersin &
<sub>6</sub> Min, 1993; Franco, Malhotra, & Simonovits, 2014; Fraser, Parker, Nakagawa, Barnett, &
<sub>7</sub> Fidler, 2018; Gerber, Green, & Nickerson, 2001; Gopalakrishna et al., 2022; John,
<sub>8</sub> Loewenstein, & Prelec, 2012; Kepes, Keener, McDaniel, & Hartman, 2022; Stefan &
<sub>9</sub> Schönbrodt, 2023). In this format, peer review takes place before data collection and the
<sub>10</sub> decision to publish is made before authors, reviewers, and editors know the study results. In
<sub>11</sub> addition to preventing editors from selectively rejecting unfavourable results (in particular
<sub>12</sub> negative or null results), this is thought to remove incentives for authors to hide, embellish,
<sub>13</sub> or misrepresent results because publication no longer depends on them (Chambers, Dienes,
<sub>14</sub> McIntosh, Rotshtein, & Willmes, 2015). Initial evidence from psychology and neighbouring
<sub>15</sub> disciplines shows that Registered Reports indeed contain much higher rates of negative
<sub>16</sub> results than the standard literature (Allen & Mehler, 2019; O'Mahony, 2023; Scheel, Schijen,
<sub>17</sub> & Lakens, 2021).

<sub>18</sub> Advocates of the format have argued that the pre-data publication guarantee should
<sub>19</sub> make Registered Reports particularly attractive to researchers (e.g., Chambers & Tzavella,
<sub>20</sub> 2022). The argument is that Registered Reports reduce uncertainty about whether and
<sub>21</sub> where a study will be published before authors have invested in conducting the study, and
<sub>22</sub> that such risk reduction is appealing in a research climate that involves substantial
<sub>23</sub> publication pressure in many countries and disciplines (Gopalakrishna et al., 2022; Miller,
<sub>24</sub> Taylor, & Bedeian, 2011; Paruzel-Czachura, Baran, & Spendel, 2021; Tijdink, Vergouwen, &
<sub>25</sub> Smulders, 2013; van Dalen, 2021; van Dalen & Henkens, 2012; Waaijer, Teelken, Wouters, &
<sub>26</sub> van der Weijden, 2018). However, if strategic concerns about publishability indeed influence
<sub>27</sub> researchers' choices for or against Registered Reports, it is unlikely that they would always

28 cause risk aversion (i.e., favouring Registered Reports as a low-risk option). Researchers'

29 willingness to take risks regarding publication success may instead vary depending on factors

30 such as available resources, time pressure, or competition. This could create situations in

31 which Registered Reports remain unpopular and would never gain traction without

32 additional incentives or interventions. And indeed, although uptake is growing exponentially

33 (Chambers & Tzavella, 2022), the market share of Registered Reports is currently still much

34 smaller than one might expect if authors saw them as unreservedly beneficial for their

35 careers. Here, we examine these possibilities with an agent-based simulation, modelling

36 authors' choices between publication formats as decision making under risk to identify

37 circumstances in which Registered Reports might be used highly selectively, or not at all.

38 **Design and intended functions of Registered Reports**

39      The review process of Registered Reports is split into two stages. At Stage 1, reviewers

40 evaluate a pre-study protocol containing the research questions, hypotheses, methods, and

41 planned analyses of a proposed study. In case of a positive decision, the journal issues an

42 'in-principle acceptance' and commits to publishing the eventual report, regardless of the

43 direction of the results. Only after in-principle acceptance, authors move on to data

44 collection and analysis and eventually complete the manuscript. At Stage 2, the final report

45 (now including the results) is subjected to a second round of peer review, but this time only

46 to ensure that the study was carried out as planned, that the data pass any pre-specified

47 quality checks, and that authors' conclusions are justified by the evidence. Manuscripts can

48 still be rejected at Stage 2, but only for substantial violations of the Stage-1 protocol or data

49 that are uninterpretable or uninformative (e.g., caused by equipment failure), not for the

50 direction or statistical significance of the results.

51      Through this process, Registered Reports address publication bias as well as so-called

52 'questionable research practices' (QRPs). These two problems are considered important

53 contributors to psychology's replication crisis (Ferguson & Heene, 2012; Wagenmakers,

⁵⁴ Wetzels, Borsboom, van der Maas, & Kievit, 2012) and to research waste in the biomedical

⁵⁵ sciences (Chalmers & Glasziou, 2009) because they skew the available evidence for scientific

⁵⁶ claims, causing overconfidence and inflated rates of false-positive inferences. Publication bias

⁵⁷ can result from editors and reviewers disproportionately rejecting submissions with negative

⁵⁸ results ('reviewer bias,' Atkinson, Furlong, & Wampold, 1982; Greenwald, 1975; Mahoney,

⁵⁹ 1977) or from researchers failing to submit negative results for publication ('file-drawering,'

⁶⁰ Ensinck & Lakens, 2023; Franco et al., 2014; Rosenthal, 1979). In Registered Reports, the

⁶¹ in-principle acceptance issued at Stage 1 addresses both of these issues: Editors and

⁶² reviewers cannot reject the Stage-2 report based on the direction of the results, which also

⁶³ reduces the incentives for authors to file-drawer the study in case of negative results. QRPs

⁶⁴ are practices that exploit undisclosed flexibility in data collection and analysis, for example

⁶⁵ when analysing different justifiable combinations of variables, subsamples, and decision

⁶⁶ criteria, and only reporting the ones with favourable results, or by presenting *post hoc*

⁶⁷ inferences as having been predicted *a priori* (Agnoli, Wicherts, Veldkamp, Albiero, &

⁶⁸ Cubelli, 2017; Fiedler & Schwarz, 2016; Fraser et al., 2018; Gopalakrishna et al., 2022; John

⁶⁹ et al., 2012; Kerr, 1998; Simmons, Nelson, & Simonsohn, 2011). Registered Reports minimise

⁷⁰ the risk of QRPs via the two-stage review process, in which the Stage-1 protocol acts as a

⁷¹ preregistration and reviewers' task during Stage-2 review is to flag any undisclosed

⁷² deviations from it.

⁷³ **Efficacy of Registered Reports**

⁷⁴ Registered Reports were first launched in 2013 at the journal *Cortex* (Chambers, 2013)

⁷⁵ and are now offered by over 300 journals, predominantly in the behavioural sciences and life

⁷⁶ sciences (see cos.io/rr). Nearly 600 Registered Reports had been published by 2021, with

⁷⁷ uptake growing exponentially (Chambers & Tzavella, 2022). In Chapter 2, we analysed the

⁷⁸ first cohort of published Registered Reports in psychology and showed that the first

⁷⁹ hypothesis reported in these articles was supported in only 44% of cases, compared to 96%

80  in a random sample of standard reports (Scheel et al., 2021). Similarly low proportions of

81  positive results were found in partially overlapping samples of Registered Reports in

82  psychology and neuroscience (39.5%, Allen & Mehler, 2019) and in psychology, neuroscience,

83  health, and education (50.1%, O'Mahony, 2023). These findings suggest that Registered

84  Reports indeed reduce biases that inflate the rate of positive results in the standard

85  literature. However, the existing estimates are based only on observational evidence and may

86  thus be confounded by other systematic differences between Registered Reports and

87  standard reports.

88      Systematic differences would act as confounders if they affected either the probability

89  of a positive result when testing a true hypothesis (statistical power) or the base rate of true

90  hypotheses. The first option is not supported by current evidence: A study comparing

91  Registered Reports with matched controls found that Registered Reports have higher median

92  sample sizes and, in blind reviews, are judged to be more rigorous in methodology and

93  analysis and of higher overall quality (Soderberg et al., 2021). Based on this finding, the

94  increased amount of negative results in Registered Reports is unlikely to be an artefact of

95  lower statistical power or poorer methods. But the second option — a difference in the rate

96  of true hypotheses, or the (prior) probability that the tested hypothesis is true — has not yet

97  been directly studied. The idea that Registered Reports might contain fewer true hypotheses

98  has some plausibility: If researchers expect that negative results are difficult to publish in

99  standard reports but pose no problem in Registered Reports, they might selectively choose

100 the Registered Report route when studying hypotheses that they think will yield negative

101 results. If researchers additionally perceive the standard publication route as less costly (e.g.,

102 more habitual, more flexible, faster, requiring lower sample sizes, etc.), standard reports

103 would plausibly remain the preferred option for hypotheses that researchers are more certain

104 are true and will yield publishable results.

105     Such an effect could explain why both Scheel et al. (2021) and Allen & Mehler (2019)

found that replication studies in the Registered Reports literature had descriptively lower rates of positive results than original studies, although the difference was not significant in either case (39% vs 50% in Scheel et al., 2021, and 34% vs 45.5% in Allen & Mehler, 2019, though note that the studied samples partially overlap). As discussed in Chapter 2, replication attempts may more often than novel research be driven by the suspicion that the tested hypothesis is not true (and that the result of the original study was a false positive). It might also partially explain differences between our results and those of O'Mahony (2023), who compared Registered Reports to standard reports that were matched on based on the publishing journal, time of publication, and to a lesser extent research topic, design, and studied population. O'Mahony finds a difference in the positive result rate of Registered Reports and standard reports half as large as the one in our study (26 vs 52 percentage points), which compared Registered Reports with a random sample of standard reports (matched only on discipline). Matching articles more closely could lead to more comparable prior probabilities of the hypotheses tested in both formats and thus account for part of this discrepancy. However, the two studies also differ in the target population and estimand (O'Mahony analysed all tested hypotheses whereas Scheel et al. focused on the first hypothesis per article), which makes the estimates difficult to compare.

This consideration highlights the importance of understanding the costs and benefits of Registered Reports from the authors' perspective. If current incentives cause Registered Reports to be used selectively in specific situations or for specific research questions, meta-scientists studying this emerging literature would need to take such factors into account. Selective use could also lead to a situation in which Registered Reports become associated with certain types of results (e.g., negative results), which could make the format unattractive in the long run if these results are deemed less interesting or important by the research community. More generally, a better understanding of the incentives driving researchers' publication choices can help determine where, when, and by whom Registered Reports are likely to be used or avoided. Such knowledge could then be used to identify

133 areas in which Registered Reports may not gain popularity naturally and anticipate the need

134 for further intervention (e.g., via policy) when there is a demand for unbiased results.

**Author incentives for Registered Reports**

136     Registered Reports are generally thought to '[neutralise] bad incentives' (Chambers,

137 2013, p. 609), in particular the incentive to exaggerate or misrepresent a study's results in

138 order to make them more publishable in the standard literature. This assumption is

139 conditioned on the format: Once authors have decided to take the Registered Report route,

140 they can improve their publication chances only via the proposed research question and

141 methods in Stage-1 review, and editors have an interest in selecting informative study

142 designs because they are bound to publishing the study's results even when they turn out

143 negative. In contrast to standard reports, the results are thus no longer a main target to

144 'hack' or select on, which should make them less biased and more trustworthy.[1]

145     The incentives for choosing the Registered Reports route in the first place, however,

146 are less clear. Advocates of the format have argued that it 'serve[s] the interests of individual

147 scientists' (p. 12, Chambers & Tzavella, 2022) because it reduces scientists' risk of investing

148 in research projects whose results turn out to be difficult to publish. The argument is based

149 on the assumptions that researchers a) are under pressure to amass journal publications

150 (which still are a central currency for hiring and promotion decisions, R. Müller, 2014; van

151 Dalen & Henkens, 2012) and b) face shortfalls in publication output when their studies yield

152 negative results (which are more difficult to publish in the standard literature due to

153 publication bias). The following quote from a talk by Chris Chambers (September 2021)

154 summarises this sentiment:

155     And the second main benefit, the one that really is the main big one, the big

---

[1] Of course authors may still be biased themselves, for example because they are hoping to find supporting evidence for their own theories. However, behaviour motivated by such biases would not increase publication chances in a Registered Report, and the two-stage review process is designed to minimise their influence on the analysis and the interpretation of results.

draw, is that as a researcher you can get your paper accepted before you even

start your research and regardless of how the results turn out in the end. So no

more playing the *p*-value lottery, gambling on certain results going a certain way,

otherwise you won't have your PhD or you won't get your next fellowship or your

next grant — takes all of that pointless, and I think quite foolish, gambling out of

the equation $(\ldots)^2$

But would researchers ever prefer to gamble? Typically, authors care not only about

their studies being published at all, but also about the reputation of the publishing journal

as well as citation rates (which are causally influenced by journal rank, Traag, 2021). In

standard reports, the career-relevant payoffs associated with a publication can thus vary

from very low, for example when authors file-drawer a manuscript because the chances of

success do not justify the cost of repeated submissions and revisions (Ensinck & Lakens,

2023), to very high, for example when a manuscript is published in an extremely high-impact

journal like *Nature* or *Science* and frequently cited. Compared to this, the payoffs from

Registered Reports have lower variance. At the lower end of the range, Registered Reports

reduce variance because they minimise the chances of a very low payoff (no publication at

all). But they also reduce variance at the upper end of the range: Upon receiving

in-principle acceptance (but before results are known), most Registered-Report authors know

that their study will not be published in a high-impact journal because very few such

journals currently offer the format. For standard-report authors, in contrast, this question is

not settled before they have analysed their data and written up the results — particularly

'good' results may offer the chance of publishing in high-impact journals (which are more

strongly affected by publication bias, Ceausu et al., 2018; Littner, Mimouni, Dollberg, &

Mandel, 2005; Mimouni, Krauthammer, Gershoni, Mimouni, & Nesher, 2015; Siontis,

Evangelou, & Ioannidis, 2011). Of course most standard reports will also not be published in

---

[2] https://youtu.be/FiVI3cwVMZI?list=PLChfyH8TVDGmYENpXUDPaeeq2SLh8q9dt&t=1047, from minute 17:27

high-impact journals, but the *chance* persists. Therefore, as long as the payoff associated

with a published Registered Report is not always on par with the best possible outcome of

the standard publication route, there will be situations in which the standard

route — 'taking the gamble' — is more beneficial for researchers.

**Publication strategies as decision making under risk**

Which are those situations? Because the payoffs of Registered Reports and the

standard publication route differ in variance, authors' choice between the two formats

represents *decision making under risk.* This framing allows us to use tools from the literature

on risk sensitivity to study when Registered Reports serve the interests of individual

scientists less well than standard reports. Decision making under risk is a broad subject with

applications in cognitive science, economics, and biology. Here, we draw on research in

behavioural ecology and use risk-sensitivity theory to model factors that influence risk

preferences and simulate their effects on researchers' publication strategies. Following

Winterhalder, Lu, & Tucker (1999), we define *risk* as 'unpredictable variation in the outcome

of a behavior, with consequences for an organism's fitness or utility' (p. 302). *Risk aversion*

thus means preferring a low-variance option over a high-variance option, and *risk proneness*

the reverse.[3] Organisms are *risk sensitive* when they are not only sensitive to the average of

outcomes of different behavioural options but also to their variance.

Risk-sensitivity theory is a normative theory developed to explain the foraging

behaviour of animals. It was originally designed to determine optimal food-acquisition

strategies for animals faced with a choice between a relatively stable (low-variance) food

source and a risky (high-variance) source that sometimes yields large payoffs and sometimes

small payoffs (or none at all). Organisms are predicted to be sensitive to such differences in

risk when payoffs (e.g., the amount of food) have non-linear consequences for the organism's

---

[3] Note that these definitions differ from the ones used in expected utility theory, where risk aversion, risk proneness, and risk indifference are defined as concave-down, convex-up, and linear utility functions, respectively.

205  survival or reproductive fitness. This is the case when, for example, additional increments of

206  food yield smaller and smaller returns for an animal's fitness, or when amounts below a

207  certain threshold would cause starvation. In psychology and economics, analogous problems

208  in human decision-making are usually studied with utility-based theories, most prominently

209  expected utility theory and prospect theory. The predictions of all three theories overlap

210  substantially, but risk-sensitivity theory uses fitness instead of utility as its central currency.

211  Despite its initially narrow scope, risk-sensitivity theory has proven itself as a powerful

212  framework for explaining risk-sensitive behaviour in a wide range of situations and species,

213  including humans (Kacelnik & Bateson, 1996; Mishra, 2014; Winterhalder et al., 1999).

**The present study**

215      In the following, we apply risk-sensitivity theory to the situation of researchers faced

216  with the choice of conducting a Registered Report or pursuing the standard publication

217  route. Using a simulation model, we explore how four aspects of academic careers and

218  incentive structures that are relevant to risk sensitivity may affect researchers' publication

219  strategies: whether additional publications yield decreasing or increasing returns for career

220  success, empirical pace (the frequency at which studies can be completed), publication

221  targets that must be met to continue or further one's career, and competition. Our goal is to

222  understand in which circumstances Registered Reports should be particularly attractive,

223  particularly unattractive, or particularly prone to selective use. The results of this analysis

224  may help anticipate research fields and career stages in which the format is unlikely to take

225  foot without additional changes to norms, incentives, or policy, and flag situations in which

226  the results of published Registered Reports may be particularly difficult to compare to the

227  normal literature. The following sections outline central concepts of risk-sensitivity theory,

228  relate them to characteristics of academic careers, and describe an evolutionary simulation

229  model in which their effects on researchers' risk-sensitive publication decisions are examined.

**Conceptual application of risk-sensitivity theory to publication decisions**

This section describes general factors that affect the role of risk for individual's fitness and connects these factors to relevant elements of academic careers. In this context, risk-sensitivity theory's focus on reproductive fitness as the central outcome may be seen as misguided. But although researchers do not forage, grow, reproduce, and die in the *biological* sense, they are undoubtedly concerned with factors that influence their survival and the propagation of their traits in an *academic* sense. As argued by Smaldino & McElreath (2016), academic research satisfies the three requirements for natural selection: variation (e.g., in research practices), consequences of this variation for survival and reproduction (e.g., some practices increase the chances of staying in academia and of being copied by others), and heritability (e.g., PhD students copying the research practices of their advisors). For natural selection to operate, we do not need to assume that researchers are consciously trying to maximise their 'academic fitness'—a competitive job market will by definition select for individuals whose past behaviour increased their prospects, regardless of their intentions. Such competition can create bottlenecks between early-career and tenured positions in many academic disciplines, which inevitably induce a selection pressure for career-promoting behaviours (Smaldino & McElreath, 2016; for a similar approach, see also Higginson & Munafò, 2016).

In applying risk-sensitivity theory to researchers' publishing behaviour, we will therefore conceptualise fitness as career success. This decision does not imply that career success is the only or the proximal motivation for researchers' behaviour in practice (or that it is a conscious motive at all), just as evolutionary theory does not imply that reproductive success is the only or the proximal motivation for human behaviour in everyday life. However, we do assume that selection for career-promoting behaviours has a noticeable impact on research practice.
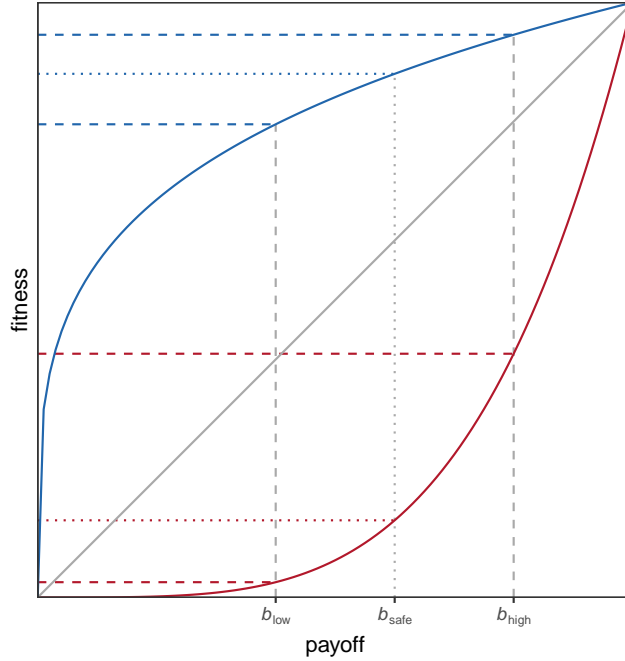
*Figure 1*. Consequences of non-linear fitness functions. Payoffs $b_{low}$, $b_{safe}$, and $b_{high}$ are converted into fitness with a diminishing (blue), linear (grey), or increasing (red) returns function.

$_{255}$ ***Non-linear fitness functions.*** The first and perhaps most ubiquitous factor

$_{256}$ leading individuals to be risk sensitive are non-linear relationships between the outcomes of

$_{257}$ an individual's behaviour (e.g., harvested food items, publications) and its reproductive

$_{258}$ success (Kacelnik & Bateson, 1997). Consider two options, $O_{safe}$ and $O_{risky}$. $O_{safe}$ always

$_{259}$ gives the same payoff $b_{safe}$, whereas $O_{risky}$ gives either a low payoff $b_{low}$ or a high payoff

$_{260}$ $b_{high}$, each with probability $\frac{1}{2}$. When $b_{safe} = \frac{(b_{low} + b_{high})}{2}$, $O_{safe}$ and $O_{risky}$ have the same

$_{261}$ expected payoff. However, we would only expect an individual to be indifferent between the

$_{262}$ two options if the consequences of their payoffs for the individual's fitness are linear. When

$_{263}$ the function relating payoffs to fitness is instead convex or concave (yielding increasing or

$_{264}$ diminishing returns, respectively), the expected fitness of $O_{safe}$ and $O_{risky}$ will differ and

$_{265}$ shift the individual's preference towards risk proneness or risk aversion. An illustration of

$_{266}$ this example is shown in Figure 1: While the payoffs $b_{low}$, $b_{safe}$, and $b_{high}$ are equidistant on

$_{267}$ the x-axis, $b_{safe}$ is associated with greater fitness than the average fitness obtained from $b_{low}$

$_{268}$ and $b_{high}$ when the function is concave, and with lower fitness when the function is convex.

In other words, $O_{safe}$ has greater expected fitness than $O_{risky}$ when returns are diminishing, and $O_{risky}$ has greater expected fitness than $O_{safe}$ when returns are increasing.

Non-linear relationships are arguably the norm in the natural world and linear relationships the exception. This plausibly holds for academia as well, where the effect of publication success on researchers' career success might change over time: For early-career researchers, small increases in the number or impact of publications may have an accelerated effect on career success, whereas any one addition to the long publication record of an established professor may have a negligible effect for their career.

***Number of decision events before evaluation.*** A second risk-relevant factor considered here is the number of decision events taking place before an individual's fitness is evaluated. When a risky option is chosen repeatedly, the average of the accumulating payoffs gets closer and closer to the long-run expected payoff. This means that the danger of loosing out completely by only acquiring the lowest possible payoff of the risky option diminishes, making the risky option relatively more attractive. However, this relationship only holds for repeated decision events *before* an individual's fitness is evaluated. When fitness is evaluated after a single decision event, a risky option is more likely to yield an extreme outcome that translates to zero fitness (i.e., death or an ultimate failure to reproduce).

In situations like this, when a single risky decision might cost an individual's life or offspring, average fitness is best described by the geometric mean instead of the arithmetic mean (Haaland, Wright, & Ratikainen, 2019). The geometric mean is more sensitive to variance because it is multiplicative, capturing the fact that one failure to reproduce can end a genetic lineage. This circumstance has been shown to produce bet-hedging: Risk-averse strategies may be more adaptive across many generations even when more risk-prone strategies produce better outcomes in any one generation, simply because risk-proneness is also more likely to lead to extinction by sheer bad luck (Haaland et al., 2019). While average fitness across generations is best represented with the geometric mean, average fitness *within*

295 a generation is better captured by the arithmetic mean, reflecting the additive accumulation

296 of payoffs from decision events before fitness is evaluated. Therefore, as the number of

297 decision events per generation (i.e., before fitness is evaluated) increases, the

298 variance-sensitive geometric mean of acquired payoffs becomes relatively less important and

299 the less variance-sensitive arithmetic mean becomes more important. Consequently, an

300 individual's behaviour should switch from relative risk-aversion to relative risk-proneness.

301      For the purpose of the present study, 'decision events' refer to researchers' decisions of

302 whether to conduct a Registered Report or pursue the standard publication route. Because

303 Registered Reports must be submitted before data collection, such decisions occur whenever

304 researchers start a new empirical project that they later may want to publish.[4] The number

305 of decision events before evaluation thus reflects the number of empirical projects that a

306 researcher can conduct before their publication record is considered for hiring, promotion, or

307 grant funding decisions. We will call this parameter 'empirical pace'.

308      Key factors influencing empirical pace are the time and resources required to conduct a

309 study and the time and resources researchers have available. Empirical pace may thus differ

310 between research areas that vary in speed and/or cost of data collection (e.g., a field relying

311 on online questionnaires *vs* a field relying on fMRI studies) or between research labs that

312 vary in funding and manpower (for a discussion of the related concept 'startup cost' and its

313 impact, see also Tiokhin, Yan, & Morgan, 2021). Even career stage might affect empirical

314 pace to some extent, for example because career progress often comes with increased funding

315 and the supervision of junior researchers whose efforts boost the supervisors' output (R.

316 Müller, 2014), and because junior researchers often have short-term contracts that limit the

317 available time for producing research output before their CVs are evaluated for the next

318 application.

———

[4] At the current moment, most researchers likely never consciously consider Registered Reports as a publication option. However, the fact that they *could* nonetheless renders their pursuit of standard publications a choice, albeit an implicit one.
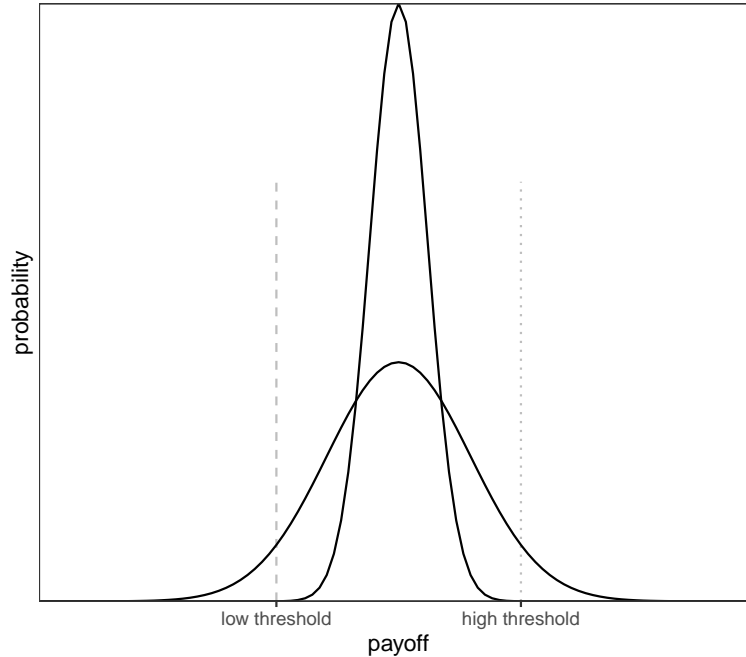
*Figure 2*. Survival thresholds. When fitness drops to zero below the low threshold (dashed line), individuals should be risk-averse because the outcomes of the low-risk option (narrow distribution) are guaranteed to lie above the threshold and the outcomes of the high-risk option (wide distribution) have a non-negligible risk of falling below the threshold. When fitness drops to zero below the high threshold (dotted line), individuals should be risk-prone because only the high-risk option provides a chance of passing the threshold.

***Survival thresholds and competition.*** A final important factor for risk-sensitive behaviour are thresholds for survival and reproduction (Hurly, 2003; Winterhalder et al., 1999). Survival thresholds are cutoff points below which an individual's fitness drops to zero, for example due to starvation. Risk-sensitivity theory predicts that an individual will be risk averse when the resources provided by a low-variance option are sufficient to meet the threshold and risk-prone when they are not (Mishra, 2014). For example, a hummingbird that needs to acquire a certain amount of calories to survive the night will prefer a low-risk food source if the expected amount of calories is above the threshold, but avoid the low-risk source if only a higher-risk source provides a chance of survival. One such situation is depicted in Figure 2.

Although comparable cutoff points in academic careers may have somewhat less severe consequences, they certainly exist: The number and impact of a researcher's publications are

often explicit criteria in decisions that are central to the individual's career, such as whether they will be awarded a PhD (Altenberger, Leischik, Vollenberg, Ehlers, & Strauss, 2024; Muijrers, 2000), whether they will receive grant funding (Simsek, de Vaan, & van de Rijt, 2024; van den Besselaar & Leydesdorff, 2009), whether they will be offered a tenure-track position (van Dijk, Manor, & Carey, 2014), or whether they will be granted tenure or promoted to full professor (Schimanski & Alperin, 2018). In some of these situations, the cutoff points are absolute and thus resemble survival thresholds in the biological sense, for example PhD regulations that determine a minimal number of peer-reviewed publications for the award of a PhD or tenure contracts that specify minimal publication targets. In other situations, the cutoff points are relative and depend on the number of eligible candidates, for example when grant funding is awarded to the 10 highest-ranked research proposals or a job is offered to the best candidate from a pool of applicants. In cases like these, one individual's success diminishes the chances of another — they represent *competition.* In the following, survival thresholds and competition will be treated as separate concepts to examine their differential effects on researchers' publication behaviour.

Each of the risk-relevant factors described above — non-linear fitness functions, empirical pace, survival thresholds, and competition — likely impacts researchers' decision strategies, including their choices between low-risk and high-risk publication options. To better understand when a low-risk option like Registered Reports should be particularly attractive or unattractive, we examine the individual and interactive effects of these factors in a simulation model.

## Simulation model

We develop an evolutionary agent-based model which simulates a population of researchers who test hypotheses, (attempt to) publish the results either as Registered Reports or as standard reports, accumulate the payoffs for successful publications, and pass their publication strategies on to the next generation of researchers.

357   ***Research phase.***   Consider a population of $n = 500$ researchers. Each researcher

358   has a fixed publication strategy $s$, the so-called submission threshold. In each round of the

359   research phase, researchers randomly choose a hypothesis to test in a study. Hypotheses are

360   true with prior probability $p$, which is uniformly distributed[5] between 0 and 1 and known to

361   the researcher. Before testing their chosen hypothesis, a researcher compares the prior $p$ of

362   their hypothesis with their publication strategy $s$. When $p < s$, the researcher chooses to

363   play it safe and conduct a Registered Report to test the hypothesis. When $p \geq s$, the

364   researcher chooses to gamble and test the hypothesis in a regular study which is then

365   submitted as a standard report.

366   For simplicity, we assume that $p$ is an ideal objective prior and that hypothesis tests

367   are free from additional sources of error. Thus, when a researcher tests hypothesis $i$, they

368   obtain a positive result with probability $p_i$ and a negative result with probability $1 - p_i$. If

369   the researcher chose to submit a Registered Report, their study is published regardless of the

370   result and the researcher receives a payoff $b_R$. However, if the researcher chose to submit a

371   standard report, they face rampant publication bias: Only positive results are publishable as

372   standard reports and yield a payoff $b_{pos} = 1$, whereas negative results are rejected or

373   file-drawered and yield no payoff, $b_{neg} = 0$. For all variations of the model tested here, we

374   assume that the payoff for a Registered Report falls between these bounds, such that

375   $b_{neg} < b_R < b_{pos}$. This assumption reflects the following considerations:

376   1. Due to publication bias in the standard literature, negative results are less valuable

377      than positive results ($b_{neg} < b_{pos}$), for example because they do not lead to a

378      publication at all, because only very low-impact journals are willing to publish them,

---

[5] Prior distributions in academic research are likely not uniform, but a realistic cross-disciplinary distribution is hard to establish. For example, molecular epidemiology may deal with predominantly false hypotheses (Wacholder, Chanock, Garcia-Closas, El Ghormli, & Rothman, 2004), whereas the social sciences may more commonly test hypotheses that are trivially true (e.g., because they are based on hidden tautologies, Wallach & Wallach, 1994). The uniform distribution thus represents a pragmatic, agnostic choice that is useful for understanding the basic mechanisms at play.

379     or because getting them published requires a lot of extra effort (e.g., via frequent

380     resubmissions following rejection or substantial revisions demanded by reviewers),

381     which diminishes the net reward.

382   2. For these same reasons, Registered Reports are on average more valuable than

383     standard reports with negative results ($b_{neg} < b_R$): Registered Reports are published

384     regardless of their results, whereas standard reports are affected by publication bias. In

385     the case of negative results, standard reports may thus not be published at all or

386     require more resubmissions or more extensive revisions than Registered Reports.

387   3. On average, standard reports with positive results are more valuable than Registered

388     Reports ($b_R < b_{pos}$), for example because many high-impact journals do not (yet) offer

389     Registered Reports, because not registering one's study *a priori* makes it easier to spin

390     the results to appear more impactful and thus increases the chances to be published in

391     a high-impact journal, or because Registered Reports may require more effort due to

392     their stricter quality criteria, lowering the net reward. While proponents of Registered

393     Reports may argue that the format has such tremendous advantages that authors'

394     resulting career benefits are superior to any alternative, this chapter is predicated on

395     the assumption that most researchers currently do not share this view. Once this

396     changes, the present investigation may happily become redundant.

397     This entire research cycle — choosing a hypothesis, choosing a publication route by

398   comparing its prior $p$ to one's publication strategy $s$, testing the hypothesis, and receiving

399   payoff $b_R$ for a Registered Report or $b_{neg}$ or $b_{pos}$ for a positive and negative standard report,

400   respectively — is repeated $m$ times.

401     ***Evaluation phase.***   At the end of the research phase, researchers' accumulated

402   publication payoffs $b_1 + b_2 + ... + b_m$ are translated into fitness $f$. Fitness is calculated with a

403   function characterised by exponent $\epsilon$, which determines the shape of the function. $\epsilon = 1$

404   yields a linear function, $0 < \epsilon < 1$ yields a concave function with diminishing returns, and

405   $\epsilon > 1$ yields a convex function with increasing returns (illustrated in Figure 1):

$$f = (\sum_{i=1}^{m} b_i)^{\epsilon} \tag{1}$$

However, two situations may cause a researcher's fitness to fall to zero even when their accumulated payoffs are non-zero. First, the sum of their payoffs may fall below an absolute survival threshold $\delta$, for example when a researcher fails to meet an agreed publication target by the time their 'tenure clock' runs out. Thus, when $\sum_{i=1}^{m} b_i < \delta$, $f = 0$. Second, the sum of their payoffs may fall below a relative threshold $\gamma$, which reflects the intensity of competition (e.g., for scarce research grants or positions). $\gamma$ is the proportion of researchers who are considered for reproduction. When $\gamma = 1$, all researchers in the population are considered for reproduction and their fitness is calculated according to Eq. 1. When $\gamma < 1$, the $(1 - \gamma) * 500$ least successful researchers receive zero fitness and cannot reproduce.[6] For example, $\gamma = 0.1$ means that only those researchers with accumulated payoffs in the top 10% of the population can reproduce, and the fitness of the remaining 90% is set to zero.

*Reproduction phase.*   Finally, the researchers in the current population retire and a new (non-overlapping) generation of researchers is created. A researcher in the new generation inherits their publication strategy $s$ from a researcher in the previous generation with the probability of the previous researcher's fitness (i.e., the new generation's publication strategies are sampled with replacement from the previous generation, probability-weighted by fitness). The new generation's publication strategies are inherited with a small amount of random noise, such that $s_{new} = s_{old} + w$, with $w \sim N(\mu = 0, \sigma = 0.01)$. Such hereditary transmission can be seen as reflecting mentorship and teaching (e.g., when established professors advise mentees to copy their strategies) or simply a generic social learning process in which successful researchers are more likely to be imitated by others (Smaldino &

---

[6] In the simulation, $\gamma$ is applied *after* fitness has been calculated. This has purely technical reasons and produces the same result as applying $\gamma$ to accumulated payoffs because all fitness functions are monotonic increasing. Applying the fitness function thus does not affect the rank order of researchers in the population.

427  McElreath, 2016; Tiokhin et al., 2021). Although this interpretation may be useful, the main

428  purpose of this aspect of the model is purely technical and not specifically intended to reflect

429  reality — it simply provides the machinery for determining which publication strategies are

430  optimal in the various situations we are investigating.

Table 1
*Parameter definitions and values*

| Parameter | Definition | Value [range] |
| --- | --- | --- |
| $n$ | population size | 500 |
| $g$ | number of generations | 250 |
| $p$ | prior probability of hypotheses | uniform [0–1] |
| $b_{neg}$ | payoff for negative standard report | 0 |
| $b_{pos}$ | payoff for positive standard report | 1 |
| $b_R$ | payoff for Registered Report | [.1, .2, …, .9] |
| $\epsilon$ | fitness function exponent | [0.2, 1, 5] |
| $m$ | research cycles per generation ('empirical pace') | [1, 2, 4, 8, 16, 32] |
| $\delta$ | survival threshold below which fitness = 0, expressed as a proportion of $m$ | [0, .25, .5, .75] |
| $\gamma$ | proportion of most successful researchers selected for reproduction (competition) | [1, .9, .5, .1, .05, .01] |

431      **Outcome variable $s$.**   We study how the evolution of researchers' publication

432  strategies $s$ is affected by the payoff for Registered Reports $b_R$ (relative to the payoffs for

433  standard reports, which are fixed at $b_{neg} = 0$ and $b_{pos} = 1$), by the shape of the fitness

434  function determined by exponent $\epsilon$, by the number of research cycles per generation $m$

435  (empirical pace), by survival threshold $\delta$, and by competition $\gamma$ (see Table 1 for an overview

436  of the model parameters and their values considered in the simulation). It is important to

437  keep in mind that a researcher's publication strategy $s$ is not an absolute decision: It

determines *how* the choice between Registered Reports and standard reports is made, not which format is chosen. As such, $s$ indicates the amount of risk a researcher is willing to take. Very low values of $s$ reflect risk proneness: The researcher prefers to gamble and chooses the standard publication route for almost all hypotheses they encounter, using the Registered Report route only for hypotheses that are virtually guaranteed to be false (and yield negative results). Very high values of $s$ reflect risk aversion: The researcher is unwilling to risk a negative result in a standard report and studies almost all hypotheses they encounter in the Registered Report format, reserving the standard publication route for hypotheses that are virtually guaranteed to be true (and yield positive results).

**Simulation approach.** We use the evolutionary mechanism of this agent-based model as a means for identifying optimal behaviour under different conditions. But this goal can also be achieved in other ways. One non-evolutionary alternative is to calculate expected fitness (i.e., the long-run average) for a wide range of $s$ and determine which strategy maximises it in each condition. A drawback of this approach is that it does not account for population dynamics and therefore cannot easily simulate the effects of competition. Because of this limitation, our study is based on the evolutionary model. However, we validate all analyses except those involving competition on the expected-fitness model and show that both models produce virtually identical results (see Appendix).

## Simulation results

The results of the simulation models will be presented in order of increasing model complexity. We start by explaining the very simple scenarios shown in Figure 3 and Figure 4. These scenarios are identical to situations discussed above and the results should thus be unsurprising. We begin with these deliberately simple situations to help readers understand the basic functioning of our model as well as the less intuitive results presented later.

When interpreting the results below, one should bear in mind that the analysed parameter values are inherently arbitrary. Although the model parameters are intended to
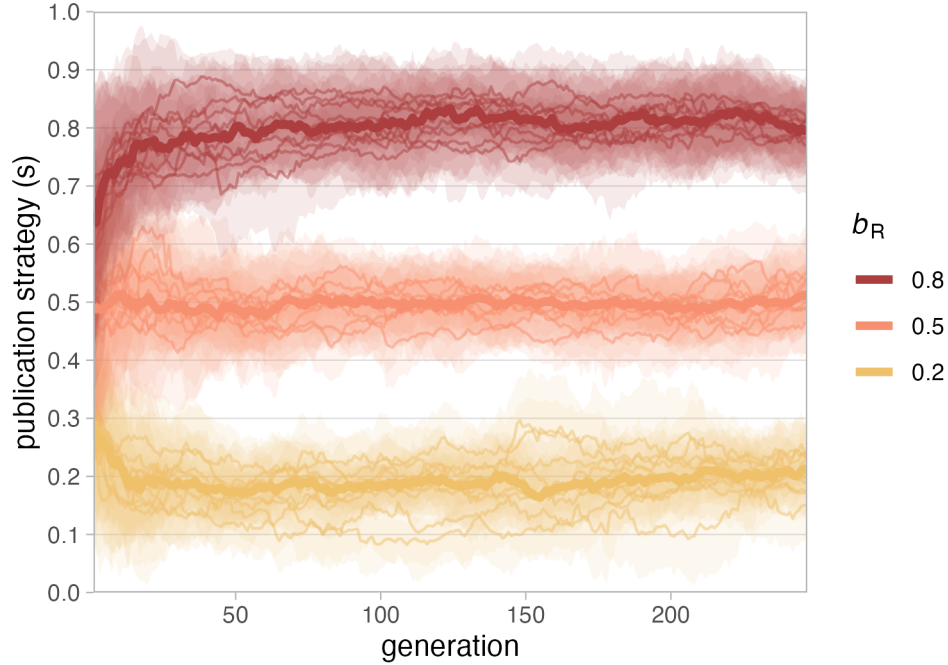
*Figure 3.* Evolution of publication strategy $s$ with 3 different payoffs for Registered Reports ($b_R$). Simulations are based on a population of $n = 500$ researchers over 250 generations, with payoffs for standard reports fixed at 0 for negative results ($b_{neg} = 0$) and 1 for positive results ($b_{pos} = 1$), a linear fitness function $\epsilon = 1$, one research cycle per generation ($m = 1$), no survival threshold ($\delta = 0$) and no competition ($\gamma = 1$). Each condition was run 10 times. Thin lines represent the median publication strategy of the population in each run, shaded areas represent the inter-quartile range of publication strategies in the population in each run, and thick lines represent the median of run medians per condition.

⁴⁶⁴ capture important characteristics of real-world concepts, their values do not represent

⁴⁶⁵ real-world units. The goal of this analysis is to understand the relative effects of the model

⁴⁶⁶ parameters in a simplified system, which means that the results are only meaningful in

⁴⁶⁷ relation to each other.

**Single research cycle per generation, linear fitness function**

⁴⁶⁹      The first generation of researchers in each simulation run is initialised with randomly

⁴⁷⁰ distributed publication strategies $s$ (drawn from a uniform distribution [0–1]), which are

⁴⁷¹ then allowed to evolve over the subsequent generations. Figure 3 shows the effect of varying

⁴⁷² the payoffs for Registered Reports when the fitness function is linear ($\epsilon = 1$), with no

⁴⁷³ survival threshold ($\delta = 0$), no competition ($\gamma = 1$), and one research cycle per generation

474  ($m = 1$). In this very simple scenario, evolved publication strategies ($s$) approximate the

475  payoff for Registered Reports in each condition, indicating that the optimal publication

476  strategy is always equal to $b_R$ ($s_{optimal} = 0.2$ when $b_R = 0.2$, $s_{optimal} = 0.5$ when $b_R = 0.5$,

477  $s_{optimal} = 0.8$ when $b_R = 0.8$). The reason behind this is the uniform distribution [0–1] of

478  hypothesis priors, the payoff structure $b_{neg} = 0$ and $b_{pos} = 1$, and the linear fitness function

479  ($\epsilon = 1$ means that fitness equals payoff). In this constellation, the expected fitness obtained

480  from a standard report is always equal to the prior of the tested hypothesis:

$$E[f_{SR}] = (p * b_{pos} + (1 - p) * b_{neg})^1 = p * 1 + (1 - p) * 0 = p \qquad (2)$$

481       For example, testing a hypothesis with $p = 0.2$ in a standard report would yield the

482  expected fitness $E[f_{SR}] = (0.2 * 1 + 0.8 * 0)^1 = 0.2$. The optimal strategy is to submit a

483  Registered Report whenever the expected fitness provided by a standard report is lower than

484  the fitness provided by a Registered Report, $E[f_{SR}] < b_R$, and thus whenever $p < b_R$. This

485  ensures that researchers always get the best of both worlds, minimising shortfalls when

486  priors are (too) low and maximising winning chances when priors are (sufficiently) high. For

487  example, $b_R = 0.5$ is larger than $E[f_{SR}]$ for all hypotheses with $p < 0.5$ but smaller than

488  $E[f_{SR}]$ for all hypotheses with $p > 0.5$. In this situation, researchers who submit Registered

489  Reports whenever $p < 0.5$ and standard reports whenever $p > 0.5$ protect themselves against

490  losing a bad bet by instead taking the fixed payoff $b_R = 0.5$, but always play a good bet and

491  thus maximise their chances of winning $b_{pos} = 1$. Every alternative is inferior in the long run

492  because researchers with $s > b_R$ lose out on increased chances of publishing a standard

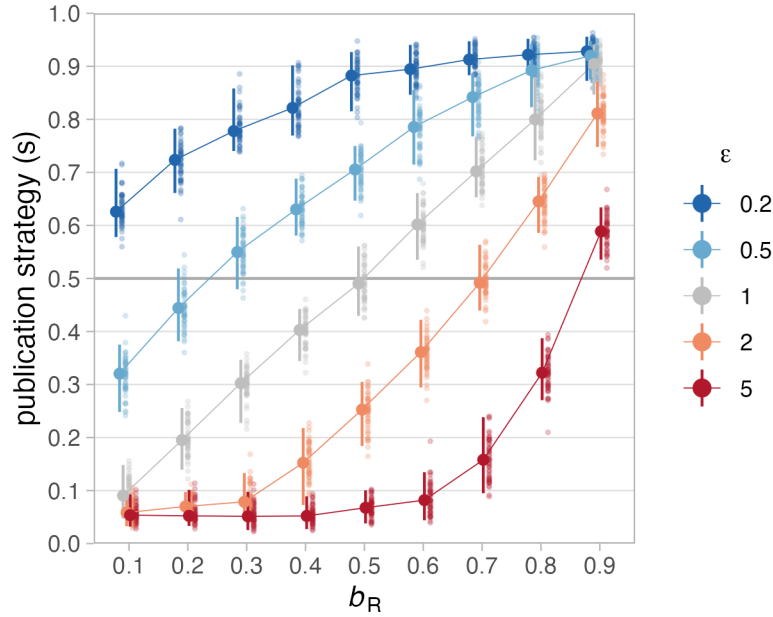493  report and researchers with $s < b_R$ take unnecessary risks and go empty-handed too often.

*Figure 4*. Effect of fitness functions on evolved publication strategies. Shown are median publication strategies in the final ($250^{th}$) generations of 50 runs for different values of $b_R$ (x-axis) and different fitness functions (characterised by exponent $\epsilon$), with one research cycle per generation ($m = 1$), no survival threshold ($\delta = 0$) and no competition ($\gamma = 1$). Fitness functions with $\epsilon = 0.2$ and $\epsilon = 0.5$ (blue lines) are concave with diminishing returns, functions with $\epsilon = 2$ and $\epsilon = 5$ (red lines) are convex with increasing returns, and the function with $\epsilon = 1$ (grey line) is linear. Small dots represent median $s$ of the final generation in each run, large dots represent the median of these 50 run medians per condition. Error bars represent the 95% capture probability around the median of medians.

## Allowing for non-linear fitness functions

Arguably, the career benefits researchers receive from publications in the real world are rarely, if ever, linear. In early career, we may assume a convex fitness function, with each addition to the short publication record of a young researcher yielding increasing returns for their prospects on the job market and their ability to obtain grant funding. A notable exception may be PhD students who plan to leave academia after their degree, and for whom the career returns of publications exceeding the PhD requirements are strongly decreasing (concave fitness function). Researchers who stay in academia may experience that the career returns for each additional publication decrease as their publication record grows, meaning that advanced career stages may also be characterised by a concave fitness function.

Figure 4 contrasts the effects of two concave and two convex fitness functions with a linear function for different payoffs for Registered Reports, in the same simple scenario with only one research cycle per generation. The grey line for $\epsilon = 1$ represents the already familiar situation from Figure 3 above: When the fitness function is linear, the optimal strategy is $s_{optimal} = b_R$. Non-linear fitness functions deviate from this pattern exactly as expected based on Figure 1. When additional payoffs yield diminishing returns ($\epsilon < 1$), Registered Reports become more attractive even when they are worth less than the expected payoff for standard reports. As explained above, this is because concave functions 'shrink' the difference between moderate and high payoffs relative to the difference between low and moderate payoffs. Conversely, when additional payoffs yield increasing returns ($\epsilon > 1$), Registered Reports are unattractive unless their payoffs are almost as large as those for published standard reports because convex functions increase the difference between moderate and high payoffs relative to low versus moderate payoffs.

When different fitness functions are taken to reflect different career stages, this pattern suggests that Registered Reports should be more attractive for senior researchers and a tough sell for early-career researchers. Interestingly, preliminary empirical evidence suggests the opposite: Registered Reports appear to be more likely to have early-career researchers as first authors than standard reports (77% vs 67% in the journal *Cortex*, Chambers & Tzavella, 2022). One explanation for this counterintuitive result could be that Registered Reports are disproportionally used by early-career researchers who intend to leave academia and thus have a concave fitness function. Alternatively, factors or dynamics not considered in this simulation may swamp out the effects of concave *vs* convex fitness functions, such as younger researchers being more likely to adopt new methods. However, as we will see below, the effects of different fitness functions are not always as straightforward as in the simple case illustrated in Figure 4 and produce different results in interaction with other risk-related factors.
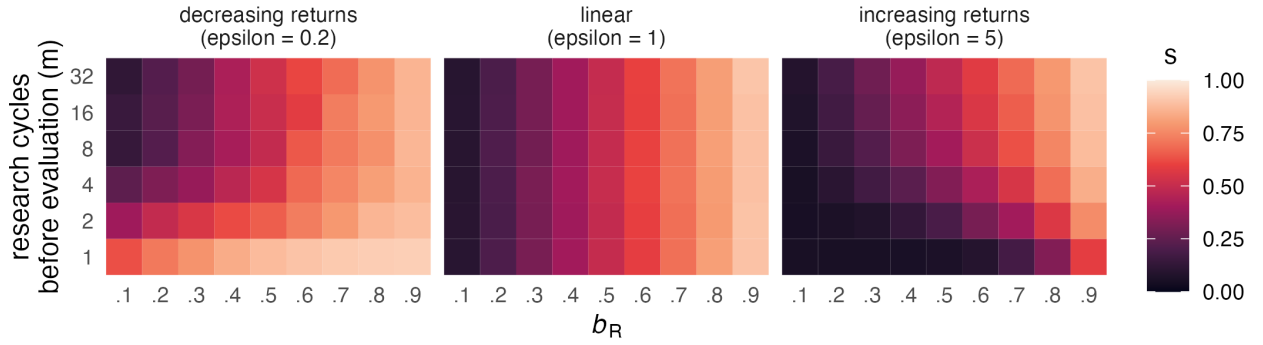
*Figure 5*. Effect of research cycles per generation (empirical pace) on evolved publication strategies. Shown are median evolved publication strategies ($s$) after 250 generations in 50 runs (tile colour represents the median of 50 run medians) depending on the number of research cycles per generation ($m$, y-axis), different values of $b_R$ (x-axis), and different fitness functions (characterised by exponent $\epsilon$) with no survival threshold ($\delta = 0$) and no competition ($\gamma = 1$).

## Varying the number of research cycles per generation (empirical pace)

The analyses presented so far focused on the simple case of one research cycle (or decision event) per generation, meaning that researchers' fitness was calculated based on the payoff from one single study. As discussed above, increasing numbers of decision events prior to evaluation may make individuals more risk-prone because single negative outcomes are less catastrophic for reproduction (Haaland et al., 2019). However, Figure 5 shows that this is not universally true — rather, the effect of increasing numbers of research cycles per generation ($m$) depends on the shape of the fitness function. Moving up on the y-axis of each panel, we see that $s$ decreases (indicating greater risk proneness) only when the fitness function is concave ($\epsilon = 0.2$, left panel) but stays constant when it is linear ($\epsilon = 1$, middle panel) and even *increases* when it is convex ($\epsilon = 5$, right panel).

Why does $m$ appear to have opposite effects for concave and convex fitness functions? As a starting point, it helps to first consider only the bottom row of each panel, where $m = 1$. These three rows contain the same results as the top, middle, and bottom curves in Figure 4 and show risk aversion when $\epsilon = 0.2$ (i.e., Registered Reports are attractive even when they yield a low payoff), risk proneness when $\epsilon = 5$ (Registered Reports are

546  unattractive even when they yield a high payoff), and a linear strategy $s_{optimal} = b_R$ when

547  $\epsilon = 1$. From this starting point, the two panels with non-linear fitness functions start to

548  approximate the linear case as $m$ increases. This pattern reflects the idea that fitness is

549  better captured by the geometric mean when $m$ is low, and better captured by the

550  arithmetic mean when $m$ is high (Haaland et al., 2019).

551      To better understand this dynamic, let's consider two researchers with extreme

552  submission strategies: Regina Register conducts only Registered Reports ($s_{Regina} = 1$),

553  Darren Daring conducts only standard reports ($s_{Darren} = 0$). The payoff for Registered

554  Reports is fixed at $b_R = 0.5$. After one research cycle, Regina receives a payoff of 0.5 and

555  Darren receives either 0 or 1 (with 50/50 odds). If fitness is calculated after this one round

556  with $\epsilon = 0.2$ (concave function, yielding diminishing returns), Regina's fitness is

557  $f_{Regina} = \frac{1}{2}^{\frac{1}{5}} = 0.87$, and Darren's fitness is either $f_{Darren-} = 0^{\frac{1}{5}} = 0$ or $f_{Darren+} = 1^{\frac{1}{5}} = 1$.

558  In a population of 100 Reginas and 100 Darrens, there will be roughly 50 lucky Darrens who

559  get a positive result and 50 Darrens who get a negative result. Lucky Darrens have a narrow

560  fitness advantage over all Reginas (1 versus 0.87), while unlucky Darrens lose to all Reginas

561  by a wide margin (0 versus 0.87). Since there are 100 Reginas but only around 50 lucky

562  Darrens, the Regina strategy is relatively more successful.

563      Let's now consider the same scenario with $m = 4$ research cycles per generation.

564  Reginas receive the same payoff in every round and accumulate $b_{total} = \frac{1}{2} * 4 = 2$. Lucky

565  Darrens (who win every time) accumulate $b_{total} = 1 * 4 = 4$, while unlucky Darrens (who lose

566  every time) again receive 0 total payoff. Now, however, the probabilistic outcomes over 4

567  rounds lead to three additional versions of Darren: moderately lucky (winning 3/4 times),

568  average (2/4, receiving the same total payoff as Reginas), and moderately unlucky (1/4).

569  Translating payoffs into fitness, the Regina strategy ($f_{Regina} = 2^{\frac{1}{5}} = 1.15$) still yields an

570  enormous advantage compared to unlucky Darrens ($f_{Darren_{unlucky}} = 0$) and only a small

571  disadvantage compared to lucky Darrens ($f_{Darren_{lucky}} = 4^{\frac{1}{5}} = 1.32$). But this time, there are

572  fewer Darrens who are less successful than Reginas because Reginas now share their place

573  with average Darrens. The relative fitness advantage of the Regina strategy thus decreases.

574  As the rate of research cycles per generation grows, the law of large numbers dictates that

575  more and more Darrens achieve average total payoffs, while fewer and fewer Darrens achieve

576  extreme total payoffs (winning 32 times in a row is much less probable than winning 4 times

577  in a row). This reduces the width of the Darren distribution until it approximates the

578  Regina distribution—meaning that optimal publication strategies become identical to those

579  optimal for a linear fitness function.

580      When the fitness function is convex ($\epsilon = 5$, yielding increasing returns), the overall

581  effect of increasing values of $m$ is the same, with the difference that Reginas are initially

582  disadvantaged (because their fitness distance to the lucky half of Darrens is much greater

583  than than to the unlucky Darrens). With larger $m$, more and more Darrens receive average

584  total payoffs and share Reginas' disadvantaged position (decreasing Regina's relative

585  disadvantage), until the Darren distribution is again virtually equal to the Regina

586  distribution. Rather than causing absolute risk aversion, increasing values of $m$ thus counter

587  the effect of $\epsilon$ and reduce the effects of concave and convex fitness functions to the linear

588  case. Consequently, the top rows ($m = 32$) of the left and right panels in Figure 5 resemble

589  the stable pattern across all $m$ shown in the middle panel.

590      Translated into terms of academic careers, this less intuitive pattern indicates that

591  being able to complete empirical studies at a higher rate—e.g., when working in a field

592  where data collection is fast and cheap or when having more resources for data collection

593  available—may cancel out the effects of different career stages. This could partly explain

594  why Registered Reports appear to be less popular among senior researchers (Chambers &

595  Tzavella, 2022) than we would expect based on the effects of different fitness functions alone:

596  Although additional publications likely yield diminishing returns in later career stages

597  (concave fitness function), academic seniority often comes with resources that boost research

<sup>598</sup> output per time (e.g., more lab members). As a consequence, established professors may be

<sup>599</sup> relatively indifferent to Registered Reports. Regarding junior researchers (for whom

<sup>600</sup> additional publications have increasing returns on career success), the results suggest that

<sup>601</sup> they may be especially reluctant to use Registered Reports when they have very limited time

<sup>602</sup> or resources to produce publications before an important selection event, such as on

<sup>603</sup> short-term postdoc contracts (R. Müller & de Rijcke, 2017).
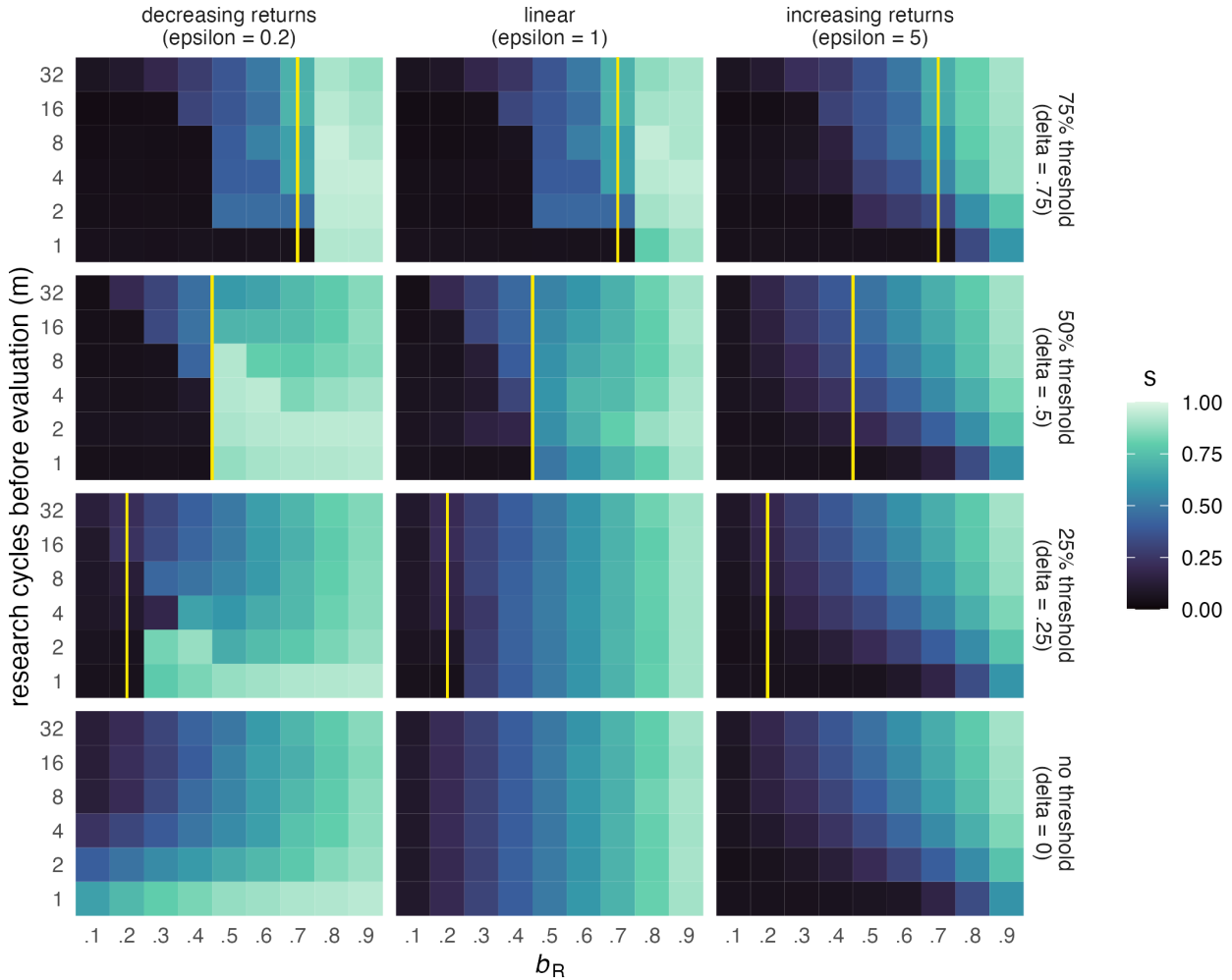


*Figure 6.* Effect of survival thresholds on evolved publication strategies. Shown are median publication strategies ($s$) after 250 generations in 50 runs (tile colour represents the median of 50 run medians) depending on survival thresholds ($\delta$, shown as vertical yellow line), fitness functions (characterised by exponent $\epsilon$), numbers of research cycles per generation ($m$), and values of $b_R$, in the absence of competition ($\gamma = 1$). Survival thresholds are set as a proportion of $m$, i.e., as a percentage of the maximum possible payoff in each condition. To reproduce, researchers must accumulate a total payoff exceeding $\delta * m$.

**Absolute survival thresholds**

The survival thresholds ($\delta$) in our model represent absolute publication targets that researchers must meet in order to progress in their career. The clearest examples for such thresholds are PhD regulations and tenure agreements. To be awarded with a PhD, many institutions and faculties require candidates to have a certain number of their thesis chapters published in peer-reviewed journals (Altenberger et al., 2024; Muijrers, 2000). Similarly, tenure agreements may include publication targets in the form of a minimum number of peer-reviewed publications within a certain time, sometimes also specifying minimal journal ranks (Liner & Sewell, 2009; Schimanski & Alperin, 2018). Such requirements may represent low, medium, or high survival thresholds depending on how demanding they are (e.g., the number of thesis chapters that must be published).

We investigate the effects of survival thresholds representing 25%, 50%, and 75% of the maximum possible payoff researchers can achieve in one generation. When $\delta > b_R$, Registered Reports alone are not sufficient to reach the survival threshold ($b_R$ values to the left of the yellow line in Figure 6). For example, at $m = 4$, a survival threshold of 75% ($\delta = .75$) means that researchers must gain at least 3 points to be able to reproduce. When $b_R = .7$, submitting four Registered Reports will only amount to 2.8 points in total, just short of meeting the threshold. On the other hand, when $b_R = .8$ (i.e., just above $\delta$), four Registered Reports would yield 3.2 points and thus ensure reproduction. Choosing the standard route some of the time can increase fitness even further, but also increases the risk of not meeting the survival threshold. As a consequence, one may intuitively expect Registered Reports to be popular whenever $\delta \leq b_R$ and unpopular whenever $\delta > b_R$.

Figure 6 shows that this is true in many, but not all conditions. First, we can see that survival thresholds have their biggest effect when the number of research cycles per generation is low — at high values of $m$, publication strategies are virtually unaffected in all conditions. Second, survival thresholds have a stronger effect when the fitness function is

630 linear ($\epsilon = 1$) or concave ($\epsilon = 0.2$). In these two conditions, they produce very similar

631 patterns: The Registered Report route is almost never chosen when $b_R$ is too low to meet

632 the survival threshold (particularly at $\delta = .25$ and $\delta = .5$; less so at $\delta = .75$), and this effect

633 tapers off as the number of research cycles increases. Compared to baseline, the change is

634 particularly striking for the concave fitness function ($\epsilon = 0.2$, left column in Fig. 6), where

635 RRs are normally preferred at low $m$. When the survival threshold is high ($\delta = .75$) or the

636 fitness function is concave, we can also see that Registered Reports become *more* popular

637 than baseline when they are worth just enough to pass the survival threshold. For the

638 convex fitness function ($\epsilon = 5$) on the other hand, survival thresholds of 25% and 50% seem

639 to have no effect at all. Only a high threshold of 75% makes RRs even less popular when

640 they have low value ($b_R \leq 0.4$), especially when the number of research cycles is low.

641       What does this mean in practice? In our model, fitness (according to the three

642 different fitness functions) is calculated after the survival threshold has been met. This is

643 meant to mimic publication requirements that are expressed in raw numbers. Importantly, it

644 also means that our simulation shows which strategies during a PhD or on the tenure track

645 lead to maximal fitness *after* researchers have successfully obtained their PhD or have been

646 granted tenure. With this in mind, it becomes easier to understand the meaning of the

647 different fitness functions. As discussed above, PhD candidates plausibly receive increasing

648 returns for additional publications (convex fitness function), unless they intend not to stay in

649 academia, in which case returns are strongly decreasing (concave fitness function). For

650 researchers on the tenure track, the fitness function after achieving tenure is also likely

651 concave, assuming a) that achieving tenure is one of the most important career goals for

652 many (making further progress relatively less important) and b) that such individuals have

653 already built up substantial publication records, to which any single addition makes less and

654 less of a difference. However, exceptions from this scenario may well exist, for example in

655 situations where tenured researchers are under great pressure to obtain grant funding.

656    Translated to real-world scenarios, our results thus suggest the following implications:

657  First, survival thresholds are almost irrelevant when researchers can complete large numbers

658  of studies before they are evaluated (reflecting characteristics of the research field, available

659  resources, or length of the evaluation period). Second, researchers with a convex fitness

660  function — such as PhD candidates who are pursuing an academic career — are only affected

661  by high survival thresholds, which lead them to choose Registered Reports even less often

662  than normal when their value is low. Third, researchers with a concave fitness

663  function — such as tenure candidates or PhD students who aim for careers outside of

664  academia — are highly sensitive to the value of Registered Reports: They virtually never

665  conduct Registered Reports when their value is too low for meeting the survival threshold,

666  but strongly prefer them when their value is sufficient (especially when empirical pace is low

667  and/or the survival threshold is high).

## Competition

669    Competition occurs whenever the demand for academic positions or grant funding

670  exceeds the supply. Figure 7 shows that competition generally leads to an avoidance of

671  Registered Reports, as can be seen by the darkening of the plots when moving up from the

672  bottom row of panels. The only exception to this rule is very low competition: When the top

673  90% are allowed to reproduce (and only the bottom 10% are rejected, $\gamma = .9$), Registered

674  Reports become more popular than they are in the absence of competition. This effect is

675  strongest for the concave fitness function ($\epsilon = 0.2$), where it holds for almost all values of $b_R$

676  at very low numbers of $m$ and for high values of $b_R$ at high numbers of $m$. When the fitness

677  function is linear or convex, Registered Reports are chosen more often only when both $b_R$

678  and $m$ are high.

679    At higher levels of competition ($\gamma > .5$), the differences between the fitness functions

680  disappear. In all three cases, Registered Reports are essentially wiped out for low numbers of
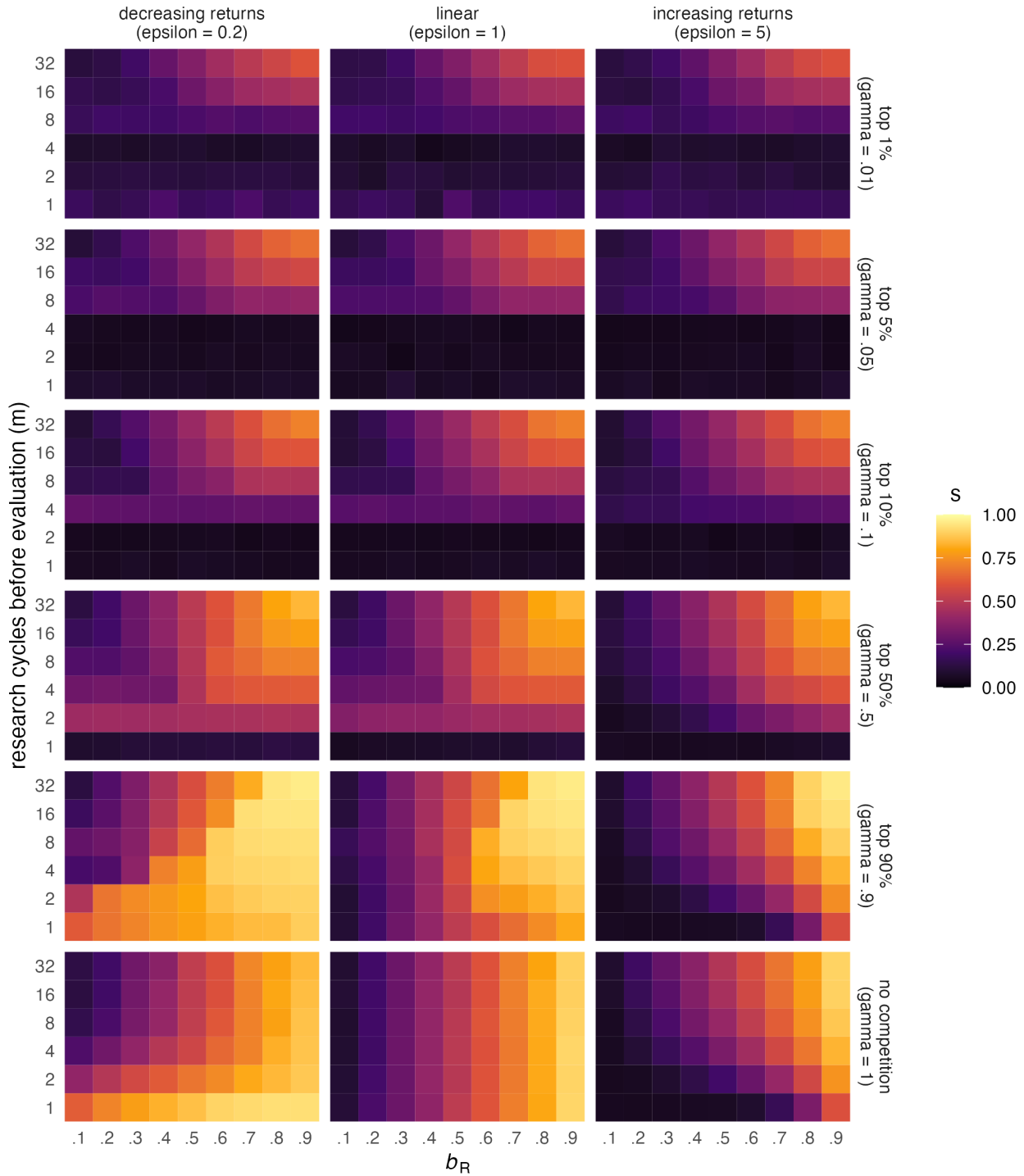
*Figure 7*. Effect of competition on evolved publication strategies. Shown are median evolved publication strategies ($s$) after 250 generations in 50 runs (tile colour represents the median of 50 run medians) depending on the intensity of competition ($\gamma$, y-axis), numbers of research cycles per generation ($m$), different values of $b_R$ (x-axis), and different fitness functions (characterised by exponent $\epsilon$) with no survival threshold ($\delta = 0$). To reproduce, researchers must accumulate a total payoff in the top $\gamma$ proportion of the population.

research cycles $(m)$,[7] and this effect increases with competition (the higher the competition,

the higher $m$ must be for Registered Reports to still be viable). Intense competition also

negatively affects Registered Reports at high numbers of $m$, but here the general pattern of

the baseline condition (a linear increase of Registered Reports popularity with $b_R$) remains

intact.

In the academic world, researchers compete for tenured positions and grants. The level

of competition may vary between research areas, countries, institutions, grant programmes,

and so on. Our findings suggest that intense competition may be a significant threat for the

viability of Registered Reports, regardless of career stage. This effect is particularly extreme

when very few research cycles can be completed before an evaluation event (e.g., in fields

working with expensive data, in labs with few resources, or on short-term contracts): In such

situations, publication strategies that involve any amount of Registered Reports are only

viable when competition is so high that success requires extraordinary luck. In contrast, very

low but non-zero levels of competition increase the popularity of Registered Reports,

especially when their value is high, when the fitness function is concave (e.g., in later career

stages), and when researchers can complete many studies before being evaluated.

---

[7] The extreme effect of competition at low $m$ appears to decrease slightly when competition is highest ($\gamma = .01$), indicated by the dark bar at the bottom of each panel becoming slightly lighter. The explanation for paradoxical result is that competition is so extreme that the natural selection in our model starts operating more on chance than on individuals' traits. Essentially, only individuals with the maximum possible payoff (publishing only standard reports with positive results) are able to reproduce. Most likely to receive this maximum payoff are individuals who investigate hypotheses with high prior probabilities. In our model, this is not a trait that can be passed on, but determined by random chance. Among individuals who experience this kind of luck, the variance of publication strategy $s$ should be high: A hypothesis with prior $p = .95$ will be submitted as a standard report and likely yield a positive result (and thus the maximum payoff) regardless of whether the researcher's publication strategy is as low as $s = .1$ or has high as $s = .9$. The higher average $s$ at low $m$ under extreme competition thus reflects relaxed selection pressure on $s$. A clearer illustration of the effect can be found in Figure A4 in the Appendix, which shows large increases in the variance of evolved publication strategies in these conditions. At higher $m$, selection on $s$ stays intact simply because much fewer individuals will be very lucky 4, 8, 16, or 32 times in a row than once or twice in a row, and publication strategy thus remains an important factor. This effect of relaxed selection is not an arbitrary feature of our model, but commonly encountered in natural populations (Snyder, Ellner, & Hooker, 2021). In many species, luck can have an outsized impact on survival and reproduction, rendering the effects of individual traits relatively less important. Luck does not eliminate natural selection, but it can significantly slow it.

## Discussion

In the artificial world of the model presented here, the standard publication route is a coin toss — the probability of obtaining a publishable result is 50% on average[8], translating to an expected payoff of 0.5 points per study. If Registered Reports are a safe alternative to this gamble and guarantee publication in every case, one might think that payoff-maximising researchers would prefer them whenever they are worth more than the expected payoff from standard reports and avoid them whenever they are worth less. This intuition, however, rests on the assumption that the career benefits researchers receive from publications are linear and involve no step changes.[9] We argue that this assumption is violated in many, if not all, real-world situations. Here, we investigated the impact of four factors that likely shape real-world situations: convex vs concave fitness functions (additional publications yielding either increasing or decreasing returns, reflecting early vs later career stages), empirical pace (reflecting differences in speed and cost of data collection, available resources, or available time), survival thresholds (reflecting absolute publication targets researchers must meet in a given time), and competition for jobs or grants. Our results show that in isolation or combined, many of these factors would lead researchers with career-maximising strategies to avoid Registered Reports — even when Registered Reports are worth more than the expected payoff from standard reports.

To summarise the results, it is useful to take the middle panel of Figure 5 ($\epsilon = 1$) as a baseline. In this panel, publication payoffs translate into linear career benefits (the fitness curve is linear and there is no survival threshold and no competition), and the outcome is highly intuitive: Researchers prefer Registered Reports whenever they are worth more than

---

[8] This is the case because we modelled the prior probability of tested hypotheses as being uniformly distributed between 0 and 1 and as being identical to the probability of obtaining a positive (i.e., publishable) result.

[9] Linearity is violated when the fitness function is concave or convex ($\epsilon \neq |1|$), but also in the presence of survival thresholds or competition, because these effectively introduce a step-change in the fitness function (low but non-zero payoffs yield zero career benefits).

719 0.5 points, and their preference is exactly proportional to $b_R$ and not affected by empirical

720 pace. Compared to this baseline, Registered Reports are *less* popular when a) additional

721 publications yield increasing returns (e.g., in early career) and empirical pace is low, b) when

722 researchers face a survival threshold that cannot be met with Registered Reports alone,

723 especially when publications yield decreasing returns once the threshold has been met (e.g.,

724 in advanced career stages) and empirical pace is low, and c) when there is substantial

725 competition. Competition has the most extreme effect and can cause a complete avoidance

726 of Registered Reports when empirical pace is low. Conversely, Registered Reports are *more*

727 popular than at baseline when a) additional publications yield decreasing returns and

728 empirical pace is low, b) Registered Reports are worth just enough to reach a survival

729 threshold and publications yield decreasing returns after the threshold is met, especially

730 when empirical pace is low, and c) when there is very low but non-zero competition,

731 especially when publications yield decreasing returns or empirical pace is high.

732      Looking at the interactions of the different factors, three observations stand out. First,

733 high empirical pace attenuates the effects of all other factors — at the highest pace we

734 considered (32 research cycles before evaluation), outcomes are identical to baseline in

735 almost all conditions. The only exception to this rule is high competition, but although

736 Registered Reports are relatively less attractive in this condition, the basic pattern is

737 preserved and they remain viable when their value is high. Second, the effect of survival

738 thresholds strongly depends on the shape of the fitness function, suggesting that publication

739 targets may have the strongest impact in advanced career stages. Third, the opposite is true

740 for high competition, which cancels out the effects of different fitness functions and thus

741 appears to have virtually the same impact across career stages.

## Implications

743      Our model predicts Registered Reports to be least popular when low empirical pace is

744 combined with intense competition or with publication targets that cannot be met with

745 Registered Reports alone. Translated to real-world academia, this suggests that fields or labs

746 in which productivity is limited by lacking resources or the cost or speed of data collection

747 (e.g., research relying on expensive or rare equipment, research on populations that are

748 difficult to access) deserve special attention. Researchers in such situations may avoid

749 Registered Reports when they must achieve publication targets that ask for a minimum

750 number of publications in high-impact journals (e.g., as part of a tenure agreement), or when

751 facing substantial competition for job positions or pressure to obtain competitive grants (e.g.,

752 if salary or research time depend on bringing in grants). When competition is high, such

753 researchers may favour standard reports even if Registered Reports are almost as valuable as

754 the best possible outcome from a standard report. Given that the last decades have seen vast

755 increases in PhD students but relatively stable numbers of tenured positions in many

756 countries (Cyranoski, Gilbert, Ledford, Nayar, & Yahia, 2011), substantial competition may

757 in fact be the default in many research fields, which could be one explanation for the

758 currently low market share of Registered Reports.

759 **Possible interventions to increase the popularity of Registered Reports.**

760 What would make Registered Reports more attractive? One answer, of course, is to change

761 the just-mentioned situational factors that make Registered Reports unpopular. However,

762 with the exception of tenure agreements and PhD regulations, these factors are difficult to

763 intervene on — competition and empirical pace cannot be changed easily, if at all. A more

764 feasible approach may be to change the payoff structure of Registered Reports relative to

765 standard reports. In the terms of our model, this could be achieved by increasing either the

766 mean or the variance of the career-relevant payoffs that authors receive from a publication.

767 For simplicity, we treated payoffs as *net* payoffs, meaning the difference between the benefits

768 and costs of each publication route. In reality, the payoffs associated with Registered

769 Reports can thus be raised by either increasing their benefits or lowering their costs (or both)

770 relative to those of standard reports. This implies three potential targets for intervention in

total: the benefits of Registered Reports, the costs of Registered Reports, and the variance

of the net payoff of Registered Reports, each relative to standard reports.

***Increasing the benefits of Registered Reports.*** The starting point of our study

was that whether and where a study is published is partly influenced by the study's results

in standard reports, but not in Registered Reports. We thus focus on the author benefits

associated with the prestige and impact of the publishing journal (assuming that these

parameters are both directly relevant for authors and causally influence citations, another

relevant parameter, Traag, 2021). These benefits could be raised if more prestigious,

high-impact journals offered the format. High-impact journals may currently be reluctant to

offer Registered Reports for fear of being forced to publish studies with uninteresting results,

which might be cited less often. Even when offering the format in principle, the same

concern may lead such journals to be prohibitively selective during Stage-1 review and reject

nearly all proposals. Perhaps as the result of such a dynamic, the journal *Nature* launched a

Registered-Reports submission track in February 2023 (Nature, 2023), but appears to have

published at most one Registered Report by August 2024.[10] In practice, journals that are

willing to participate in raising the value of Registered Reports should thus strive for

designing an editorial process which, if ambitious, does not set unrealistic standards.

The value of Registered Reports can also be raised by those who ultimately provide the

'career-relevant benefits' associated with a publication, namely faculty committees

responsible for hiring and promotion decisions. Placing a premium on Registered Reports in

tenure agreements, promotion criteria, and hiring processes could increase the attractiveness

of the format substantially. This idea is in line with recent calls for greater emphasis on

rigorous and transparent research methods in hiring, promotion, and tenure decisions (e.g.,

---

[10] We used the search function on *Nature*'s website to search for the string 'Registered Report' (entered without quotes) in research articles published since 22$^{nd}$ February 2023 and then searched the full texts of the 72 search hits for the string 'registered'. None of the articles was unambiguously marked as a Registered Report. Only one article (Aslett et al., 2024) contained the term 'registered report' and was phrased in a way that may be consistent with the Registered Reports format.

794 Moher et al., 2018), for example by including so-called 'open-science statements' in job ads

795 (Schönbrodt, 2016; Schönbrodt et al., 2018). However, most such statements currently either

796 do not define specific practices or mention only preregistration and not Registered Reports

797 (Schönbrodt et al., 2018). Explicitly highlighting Registered Reports in job ads and

798 weighting them more heavily than (even preregistered) standard reports in hiring, promotion,

799 and tenure decisions could therefore be a promising strategy.

800 ***Decreasing the costs of Registered Reports.*** Compared to standard reports,

801 Registered Reports may be more costly for authors due to the additional stage of peer review

802 and stricter requirements for methodological rigour and sample size. For example, Registered

803 Reports (but not standard reports) in *Nature Human Behaviour* must currently provide

804 sampling plans aiming for at least 95% statistical power or a Bayes factor of 10 (Nature

805 Human Behaviour, n.d.). However, although it may be relatively easy to lower such

806 standards, doing so would also lower the quality of published Registered Reports and thus

807 partly defeat their purpose of providing high-quality evidence. This problem illustrates that

808 many of the additional costs associated with Registered Reports may be 'good costs' that

809 increase the quality of the resulting publications (see also Tiokhin et al., 2021). To preserve

810 this quality, cost cutting may need to be confined to removing unnecessary inefficiencies,

811 such as certain bureaucratic aspects of the submission or review process.

812 Alternatively, the relative costs of Registered Reports could be decreased by increasing

813 the costs of standard reports. Going back to the example above, a high-impact journal like

814 *Nature Human Behaviour* could reasonably demand the same level of methodological rigour

815 from standard reports as from Registered Reports. This would reduce the marginal

816 advantage of standard reports over Registered Reports in terms of the investment required

817 from authors (making Registered Reports relatively more attractive), while at the same time

818 raising the quality of all studies published by the journal.

819      ***Increasing payoff variance.***    In the classic Registered Reports model, authors

820  must choose a journal before having full knowledge of the value of the eventual study (i.e.,

821  before results are known and the final manuscript is written up). From the authors'

822  perspective, the pre-data publication guarantee by the chosen journal puts a cap not only on

823  the worst possible outcome, but also on the best possible outcome.[11] Another approach to

824  making Registered Reports more attractive is therefore to remove the upper cap and give

825  authors more publication options after the research has been completed. This could be made

826  possible by a recent initiative: In April 2021, the post-publication peer review platform *Peer*

827  *Community In* (PCI) introduced a new model of Registered Reports in which authors are no

828  longer tied to a specific journal. *PCI Registered Reports* offers authors the regular process of

829  Stage-1 and Stage-2 review (including in-principle acceptance after Stage 1), but the end

830  result of a successful submission is simply a preprint with a so-called 'recommendation' from

831  *PCI*. Authors can subsequently publish their manuscript in one of several journals which

832  partnered with *PCI* and either rely on the *PCI* review process alone or offer a streamlined

833  review process for *PCI*-recommended preprints. Alternatively, authors are free to submit to

834  any other journal as if their manuscript were a standard report. This innovation gives

835  Registered-Reports authors significantly more freedom to capitalise on the results of their

836  study because a submission to PCI Registered Reports does not preclude the chance of a

837  high-impact publication. As of August 2024, there are 35 journals which accept Registered

838  Reports recommended by PCI without further review. With the growth of this list, and

839  particularly the inclusion of more high-impact journals, the *PCI Registered Reports* model

840  has the potential to change the incentive structure of Registered Reports in a profound way.

---

[11] In principle, authors are free to withdraw a Registered Report before publication and submit their manuscript elsewhere, but this strategy would incur additional costs and risks (a new review process with unknown outcome) and may be perceived as violating a social norm.

**Limitations and future directions**

By design, our model is based on assumptions that simplify and exaggerate some aspects of real-world academia and ignore many others. First, we use an extreme, one-dimensional concept of publication bias: All positive results are published, all negative results remain unpublished, and results are determined only by the prior probability of the hypotheses. Real-world publication decisions are of course based on many other factors as well, such as the relevance of the research question and the validity of the study design. And unlike in our model, tests of hypotheses with higher priors will not always be more publishable, for the simple reason that positive results of trivial (or previously tested) hypotheses are usually not highly valued (although it has been argued that research in social psychology is sometimes based on hidden tautologies, Wallach & Wallach, 1994).

Another approach is to model publication bias as favouring results that shift prior beliefs (Gross & Bergstrom, 2021). Adapting the model presented here to capture this concept of bias could be an interesting future direction. However, the present version of the model allows a conservative interpretation in which the prior probability of hypotheses simply reflects authors' predictions of the eventual publication value of different research questions. This interpretation is still congruent with Registered Reports and standard reports differing in risk, because publication value depends more strongly on the study results in standard reports than in Registered Reports.

Second, we assume that authors have perfect knowledge of the probability that they will obtain positive (or publishable) results. The assumption that authors have *some* prior knowledge of the results they might obtain is the starting point of our study, because this would enable strategic decisions about when to (not) use Registered Reports. As long as this assumption holds (i.e., authors are not completely ignorant), adding noise and even bias to authors' prior beliefs would have a diluting effect on the simulation results, but should not change the general pattern. Things may get more complicated, however, when considering

individual differences in prediction accuracy or bias. In our model, researchers who are better at predicting the results (or publishability) of their studies would outperform researchers whose predictions are more noisy or biased. In reality, certain biases may actually be beneficial, for example if overconfident individuals are also better at convincing editors and reviewers of the value of their studies.

A third, related limitation is that although researchers in the model know the prior of their hypotheses, they have no control over which hypotheses they test (hypotheses are randomly allocated). Of course real researchers can choose their own research questions, and this freedom may influence their publication strategies. In particular, researchers who are better at choosing research questions that are likely to result in high-impact publications (through talent or experience) may find Registered Reports less attractive. This is an example of ability-based risk taking: Individuals with traits or abilities that increase their expected payoff from a risky option[12] should be more risk-prone (Barclay, Mishra, & Sparks, 2018). A more nefarious version of this idea is that Registered Reports may be relatively unpopular among researchers who are more inclined to using questionable research practices (or even fraud) to obtain publishable or impactful results.

Fourth, we make the simplifying assumption that researchers work alone. Of course this is not true in most scientific disciplines, where team work is the default and most publications have more than one author. As a consequence, publication decisions are usually made jointly by researchers who may have different career-related needs. For example, senior researchers may often take the needs of their PhD students into account, which could lead them to behave more in line with a convex fitness function (increasing returns). This does not invalidate our results, but it means that real-world publication strategies can be mixtures of the individual strategies represented by our model. An interesting related consideration is

---

[12] This includes traits that increase the chances of winning, traits that increase the payoff when winning, or traits that buffer the impact of losses.

that researchers may be able to compensate for low empirical pace by forming larger teams, essentially sharing the payoffs from a limited number of research projects with more colleagues. Such an effect could cause publication strategies in fields with very slow and/or costly data collection to resemble those expected under higher empirical pace.

Finally, our model ignores the factor time. A common reservation towards Registered Reports is the concern that they take longer to publish because authors must wait for the outcome of the Stage-1 review process before starting the data collection. Standard reports, on the other hand, may occasionally have even longer publication delays, for example when they are rejected at several journals or when reviewers demand additional studies to be run. It is thus plausible that the formats differ in mean and/or variance of publication delays, and such differences could affect researchers' behaviour. Because humans tend to discount delayed rewards (Odum et al., 2020), researchers who believe the standard publication route to be faster may have a stronger preference for it than predicted by our model. To further investigate this possibility, data on the distribution of publication delays (from the beginning of a research project until publication) of Registered Reports and standard reports, as well as on researchers' beliefs regarding these delays, would be highly valuable.

## Conclusion

The basic mechanism underlying Registered Reports — publication decisions before results are known — is currently the most convincing proposal for curbing publication bias. By selecting studies based only on the strength of the research question and methods, Registered Reports are indeed 'aligning what is beneficial for individual scientists with what is beneficial for science' (Chambers & Tzavella, 2022, p. 29). However, the incentives for choosing the Registered Reports format in the first place may be less aligned with the interests of scientists. In this study, we examined the consequences of the pre-data publication guarantee in Registered Reports, which makes them a low-risk option compared to the standard publication route because it reduces the variance of publication outcomes.

Our results show that this feature does not make the format universally more attractive. Instead, many common situations in the academic ecosystem, such as publication targets and competition for tenured positions, may promote risk-prone publication strategies and lead researchers to avoid Registered Reports.

This suggests that the spread of Registered Reports to larger parts of the scientific literature in more disciplines is not simply a matter of time. In psychology, where the format may be best known, a generous estimate puts the annual rate of published Registered Reports at less than 0.1% of the literature.[13] This figure is much lower than the estimated prevalence of preregistration (7% in 2022, Hardwicke et al., 2024), a reform that was introduced around the same time with similar goals. Increasing the uptake of the format may require additional interventions, such as placing greater value on Registered Reports in hiring, promotion, and tenure decisions, raising the methodological requirements for standard reports in journals that offer both formats, or supporting the journal-independent model of *PCI Registered Reports* and encouraging more high-impact journals to subscribe to it. To the extent that scientific communities or external stakeholders have a demand for the kind of low-bias, high-quality evidence that Registered Reports can offer, such measures may be a worthwhile investment.

**Disclosures**

**Data, materials, and online resources.** Simulated data, code for reproducing all simulations and figures, and the Appendix are available at github.com/amscheel/rr-model. This manuscript was created using RStudio (RStudio 2022.07.2+576, RStudio Team, 2019) and R (Version 4.2.1; R Core Team, 2019) and the R-packages *bookdown* (Version 0.34; Xie, 2016), *data.table* (Version 1.15.4; Barrett et al., 2024), *ggplot2* (Version 3.5.0; Wickham,

---

[13] We estimate 600,000 journal publications per year in psychology and assume that less than 600 of them are Registered Reports (a generous estimate based on the 591 Registered Reports that had been published across disciplines by 2021, Chambers & Tzavella, 2022). The estimate of total publications per year was obtained via https://lens.org by applying the filters `Year Published = 2020--2023`, `Publication Type = journal article`, and `Field of Study = Psychology`, and dividing the 2,448,670 resulting hits by 4.

2016), *here* (Version 1.0.1; K. Müller, 2017), *knitr* (Version 1.46; Xie, 2015), *papaja* (Version 0.1.1.9001; Aust & Barth, 2018), *rmarkdown* (Version 2.26; Xie, Allaire, & Grolemund, 2018), *stringr* (Version 1.5.1; Wickham, 2023), and *tinylabels* (Version 0.2.3; Barth, 2022).

# References

Agnoli, F., Wicherts, J. M., Veldkamp, C. L. S., Albiero, P., & Cubelli, R. (2017). Questionable research practices among italian research psychologists. *PLOS ONE*, *12*(3), e0172792. https://doi.org/10.1371/journal.pone.0172792

Allen, C., & Mehler, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLOS Biology*, *17*(5), e3000246. https://doi.org/10.1371/journal.pbio.3000246

Altenberger, S., Leischik, R., Vollenberg, R., Ehlers, J. P., & Strauss, M. (2024). A comparative analysis of the doctoral regulations at the medical faculties in Germany. *International Journal of Medical Sciences*, *21*(4), 732–741. https://doi.org/10.7150/ijms.92167

Aslett, K., Sanderson, Z., Godel, W., Persily, N., Nagler, J., & Tucker, J. A. (2024). Online searches to evaluate misinformation can increase its perceived veracity. *Nature*, *625*(7995), 548–556. https://doi.org/10.1038/s41586-023-06883-y

Atkinson, D. R., Furlong, M. J., & Wampold, B. E. (1982). Statistical significance, reviewer evaluations, and the scientific process: Is there a (statistically) significant relationship? *Journal of Counseling Psychology*, *29*(2), 189–194. https://doi.org/10.1037/0022-0167.29.2.189

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown.*

Barclay, P., Mishra, S., & Sparks, A. M. (2018). State-dependent risk-taking. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1881), 20180180. https://doi.org/10.1098/rspb.2018.0180

Barrett, T., Dowle, M., Srinivasan, A., Gorecki, J., Chirico, M., & Hocking, T. (2024). *Data.table: Extension of 'data.frame'*. Retrieved from https://CRAN.R-project.org/package=data.table

Barth, M. (2022). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Ceausu, S., Borda-de-Água, L., Merckx, T., Sossai, E., Sapage, M., Miranda, M., & Pereira, H. M. (2018). *High impact journals in ecology cover proportionally more statistically significant findings*. bioRxiv. https://doi.org/10.1101/311068

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *The Lancet, 374*(9683), 86–89. https://doi.org/10.1016/S0140-6736(09)60329-9

Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex, 49*, 606–610. https://doi.org/10.1016/j.cortex.2012.12.016

Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered Reports: Realigning incentives in scientific publishing. *Cortex, 66*, 1–2. https://doi.org/10.1016/j.cortex.2015.03.022

Chambers, C. D., & Tzavella, L. (2022). The past, present and future of Registered Reports. *Nature Human Behaviour, 6*(1), 29–42. https://doi.org/10.1038/s41562-021-01193-7

Cyranoski, D., Gilbert, N., Ledford, H., Nayar, A., & Yahia, M. (2011). Education: The PhD factory. *Nature, 472*(7343), 276–279. https://doi.org/10.1038/472276a

de Vries, Y. A., Roest, A. M., Jonge, P. de, Cuijpers, P., Munafò, M. R., & Bastiaansen, J. A. (2018). The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: The case of depression. *Psychological Medicine, 48*(15), 2453–2455. https://doi.org/10.1017/S0033291718001873

Dickersin, K., & Min, Y. I. (1993). Publication bias: The problem that won't go away. *Annals of the New York Academy of Sciences, 703*, 135-146; discussion 146-148.

https://doi.org/10.1111/j.1749-6632.1993.tb26343.x

Ensinck, E., & Lakens, D. (2023). *An Inception Cohort Study Quantifying How Many Registered Studies are Published.* PsyArXiv. https://doi.org/10.31234/osf.io/5hkjz

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Fiedler, K., & Schwarz, N. (2016). Questionable Research Practices Revisited. *Social Psychological and Personality Science*, *7*(1), 45–52. https://doi.org/10.1177/1948550615612150

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Fraser, H., Parker, T., Nakagawa, S., Barnett, A., & Fidler, F. (2018). Questionable research practices in ecology and evolution. *PLOS ONE*, *13*(7), e0200303. https://doi.org/10.1371/journal.pone.0200303

Gerber, A. S., Green, D. P., & Nickerson, D. (2001). Testing for Publication Bias in Political Science. *Political Analysis*, *9*(4), 385–392. https://doi.org/10.1093/oxfordjournals.pan.a004877

Gopalakrishna, G., Wicherts, J. M., Vink, G., Stoop, I., Akker, O. R. van den, Riet, G. ter, & Bouter, L. M. (2022). *Prevalence of responsible research practices among academics in The Netherlands. 11*(471). https://doi.org/10.12688/f1000research.110664.2

Greenwald, A. G. (1975). Consequences of Prejudice Against the Null Hypothesis. *Psychological Bulletin*, *82*(1), 1–20.

Gross, K., & Bergstrom, C. T. (2021). Why ex post peer review encourages high-risk research while ex ante review discourages it. *Proceedings of the National Academy of Sciences*, *118*(51). https://doi.org/10.1073/pnas.2111615118

Haaland, T. R., Wright, J., & Ratikainen, I. I. (2019). Bet-hedging across generations can affect the evolution of variance-sensitive strategies within generations. *Proceedings of the Royal Society B.* https://doi.org/10.1098/rspb.2019.2070

Hardwicke, T. E., Thibault, R. T., Clarke, B., Moodie, N., Crüwell, S., Schiavone, S. R., . . . Vazire, S. (2024). *Prevalence of transparent research practices in psychology: A cross-sectional study of empirical articles published in 2022.* OSF. https://doi.org/10.31234/osf.io/t2zs9

Higginson, A. D., & Munafò, M. R. (2016). Current Incentives for Scientists Lead to Underpowered Studies with Erroneous Conclusions. *PLOS Biology, 14*(11), e2000995. https://doi.org/10.1371/journal.pbio.2000995

Hurly, A. T. (2003). The twin threshold model: Risk-intermediate foraging by rufous hummingbirds, Selasphorus rufus. *Animal Behaviour, 66*(4), 751–761. https://doi.org/10.1006/anbe.2003.2278

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science, 23*(5), 524–532. https://doi.org/10.1177/0956797611430953

Kacelnik, A., & Bateson, M. (1996). Risky Theories—The Effects of Variance on Foraging Decisions. *Integrative and Comparative Biology, 36*(4), 402–434. https://doi.org/10.1093/icb/36.4.402

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: Crossroads for theories of decision-making. *Trends in Cognitive Sciences, 1*(8), 304–309. https://doi.org/10.1016/s1364-6613(97)01093-0

Kepes, S., Keener, S. K., McDaniel, M. A., & Hartman, N. S. (2022). Questionable research practices among researchers in the most research-productive management programs. *Journal of Organizational Behavior, 43*(7), 1190–1208. https://doi.org/10.1002/job.2623

Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality*

*and Social Psychology Review*, *2*(3), 196–217.

https://doi.org/10.1207/s15327957pspr0203_4

Liner, G. H., & Sewell, E. (2009). Research requirements for promotion and tenure at PhD granting departments of economics. *Applied Economics Letters*. https://doi.org/10.1080/13504850701221998

Littner, Y., Mimouni, F. B., Dollberg, S., & Mandel, D. (2005). Negative Results and Impact Factor: A Lesson From Neonatology. *Archives of Pediatrics & Adolescent Medicine*, *159*(11), 1036–1037. https://doi.org/10.1001/archpedi.159.11.1036

Mahoney, M. J. (1977). Publication Prejudices: An Experimental Study of Confirmatory Bias in the Peer Review System. *Cognitive Therapy and Research*, *1*(2), 161–175. https://doi.org/10.1007/BF01173636

Miller, A. N., Taylor, S. G., & Bedeian, A. G. (2011). Publish or perish: Academic life as management faculty live it. *Career Development International*, *16*(5), 422–445. https://doi.org/10.1108/13620431111167751

Mimouni, M., Krauthammer, M., Gershoni, A., Mimouni, F., & Nesher, R. (2015). Positive Results Bias and Impact Factor in Ophthalmology. *Current Eye Research*, *40*(8), 858–861. https://doi.org/10.3109/02713683.2014.957777

Mishra, S. (2014). Decision-Making Under Risk: Integrating Perspectives From Biology, Economics, and Psychology. *Personality and Social Psychology Review*, *18*(3), 280–307. https://doi.org/10.1177/1088868314530517

Moher, D., Naudet, F., Cristea, I. A., Miedema, F., Ioannidis, J. P. A., & Goodman, S. N. (2018). Assessing scientists for hiring, promotion, and tenure. *PLOS Biology*, *16*(3), e2004089. https://doi.org/10.1371/journal.pbio.2004089

Muijrers, J. (2000). Same degree, same effort? A patchwork of differing PhD requirements throughout Europe disadvantages graduate students and compromises the quality of science. *EMBO Reports*, *1*(6), 463–464. https://doi.org/10.1093/embo-reports/kvd119

Müller, K. (2017). *Here: A simpler way to find your files.*

Müller, R. (2014). Postdoctoral Life Scientists and Supervision Work in the Contemporary University: A Case Study of Changes in the Cultural Norms of Science. *Minerva*, *52*(3), 329–349. https://doi.org/10.1007/s11024-014-9257-y

Müller, R., & de Rijcke, S. (2017). Thinking with indicators. Exploring the epistemic impacts of academic performance indicators in the life sciences. *Research Evaluation*, *26*(3), 157–168. https://doi.org/10.1093/reseval/rvx023

Nature. (2023). Nature welcomes Registered Reports. *Nature*, *614*(7949), 594–594. https://doi.org/10.1038/d41586-023-00506-2

Nature Human Behaviour. (n.d.). *Registered Reports*. https://www.nature.com/nathumbehav/submission-guidelines/registeredreports.

O'Mahony, A. (2023). *Comparative analysis of Registered Reports and the standard research literature* (PhD thesis). Cardiff University.

Odum, A. L., Becker, R. J., Haynes, J. M., Galizio, A., Frye, C. C. J., Downey, H., . . . Perez, D. M. (2020). Delay discounting of different outcomes: Review and theory. *Journal of the Experimental Analysis of Behavior*, *113*(3), 657–679. https://doi.org/10.1002/jeab.589

Paruzel-Czachura, M., Baran, L., & Spendel, Z. (2021). Publish or be ethical? Publishing pressure and scientific misconduct in research. *Research Ethics*, *17*(3), 375–397. https://doi.org/10.1177/1747016120980562

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rosenthal, R. (1979). The "File Drawer Problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

RStudio Team. (2019). *RStudio: Integrated development environment for r*. Boston, MA: RStudio, Inc.

Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in*

*Methods and Practices in Psychological Science*, *4*(2), 251524592110074.

https://doi.org/10.1177/25152459211007467

Schimanski, L. A., & Alperin, J. P. (2018). The evaluation of scholarship in academic

promotion and tenure processes: Past, present, and future. *F1000Research*, *7*, 1605.

https://doi.org/10.12688/f1000research.16493.1

Schönbrodt, F. D. (2016). *Changing hiring practices towards research transparency: The*

*first open science statement in a professorship advertisement*. Open Science

Framework.

Schönbrodt, F. D., Schramm, L. F. F., Etzel, F. T., Bergmann, C., Mellor, D. T.,

Schettino, A., . . . Wiehr, M. (2018). *Academic job offers that mentioned open science*.

OSF. https://doi.org/10.17605/OSF.IO/7JBNT

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology:

Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as

Significant. *Psychological Science*, *22*(11), 1359–1366.

https://doi.org/10.1177/0956797611417632

Simsek, M., de Vaan, M., & van de Rijt, A. (2024). Do grant proposal texts matter for

funding decisions? A field experiment. *Scientometrics*, *129*(5), 2521–2532.

https://doi.org/10.1007/s11192-024-04968-7

Siontis, K. C., Evangelou, E., & Ioannidis, J. P. (2011). Magnitude of effects in clinical

trials published in high-impact general medical journals. *International Journal of*

*Epidemiology*, *40*(5), 1280–1291. https://doi.org/10.1093/ije/dyr095

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Royal*

*Society Open Science*, *3*, 160384. https://doi.org/10.1098/rsos.160384

Snyder, R. E., Ellner, S. P., & Hooker, G. (2021). Time and Chance: Using Age

Partitioning to Understand How Luck Drives Variation in Reproductive Success. *The*

*American Naturalist*, *197*(4), E110–E128. https://doi.org/10.1086/712874

Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J., Thorn, F. S., Vazire,

S., . . . Nosek, B. A. (2021). Initial evidence of research quality of registered reports compared with the standard publishing model. *Nature Human Behaviour*, *5*(8), 990–997. https://doi.org/10.1038/s41562-021-01142-4

Stefan, A. M., & Schönbrodt, F. D. (2023). Big little lies: A compendium and simulation of p-hacking strategies. *Royal Society Open Science*, *10*(2), 220346. https://doi.org/10.1098/rsos.220346

Tijdink, J. K., Vergouwen, A. C. M., & Smulders, Y. M. (2013). Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. *PLOS ONE*, *8*(9), e73381. https://doi.org/10.1371/journal.pone.0073381

Tiokhin, L., Yan, M., & Morgan, T. J. H. (2021). Competition for priority harms the reliability of science, but reforms can help. *Nature Human Behaviour*, *5*(7), 857–867. https://doi.org/10.1038/s41562-020-01040-1

Traag, V. A. (2021). Inferring the causal effect of journals on citations. *Quantitative Science Studies*, *2*(2), 496–504. https://doi.org/10.1162/qss_a_00128

van Dalen, H. P. (2021). How the publish-or-perish principle divides a science: The case of economists. *Scientometrics*, *126*(2), 1675–1694. https://doi.org/10.1007/s11192-020-03786-x

van Dalen, H. P., & Henkens, K. (2012). Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. *Journal of the American Society for Information Science and Technology*, *63*(7), 1282–1293. https://doi.org/10.1002/asi.22636

van den Besselaar, P., & Leydesdorff, L. (2009). Past performance, peer review and project selection: A case study in the social and behavioral sciences. *Research Evaluation*, *18*(4), 273–288. https://doi.org/10.3152/095820209X475360

van Dijk, D., Manor, O., & Carey, L. B. (2014). Publication metrics and success on the academic job market. *Current Biology*, *24*(11), R516–R517. https://doi.org/10.1016/j.cub.2014.04.039

Waaijer, C. J. F., Teelken, C., Wouters, P. F., & van der Weijden, I. C. M. (2018). Competition in Science: Links Between Publication Pressure, Grant Pressure and the Academic Job Market. *Higher Education Policy*, *31*(2), 225–243. https://doi.org/10.1057/s41307-017-0051-y

Wacholder, S., Chanock, S., Garcia-Closas, M., El Ghormli, L., & Rothman, N. (2004). Assessing the Probability That a Positive Report is False: An Approach for Molecular Epidemiology Studies. *JNCI Journal of the National Cancer Institute*, *96*(6), 434–442. https://doi.org/10.1093/jnci/djh075

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

Wallach, L., & Wallach, M. A. (1994). Gergen versus the mainstream: Are hypotheses in social psychology subject to empirical test? *Journal of Personality and Social Psychology*, *67*(2), 233–242. https://doi.org/10.1037/0022-3514.67.2.233

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis.* Springer-Verlag New York.

Wickham, H. (2023). *Stringr: Simple, consistent wrappers for common string operations.* Retrieved from https://CRAN.R-project.org/package=stringr

Winterhalder, B., Lu, F., & Tucker, B. (1999). Risk-senstive adaptive tactics: Models and evidence from subsistence studies in biology and anthropology. *Journal of Archaeological Research*, *7*(4), 301–348. https://doi.org/10.1007/BF02446047

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida: Chapman and Hall/CRC.

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with R markdown.* Boca Raton, Florida: Chapman and Hall/CRC.

Xie, Y., Allaire, J. J., & Grolemund, G. (2018). *R markdown: The definitive guide.* Boca Raton, Florida: Chapman and Hall/CRC.