

Review

Using large, publicly available data sets to study adolescent development: opportunities and challenges

Rogier A. Kievit^{2,3}, Ethan M. McCormick^{2,3,a},
Delia Fuhrmann^{1,4,a}, Marie K. Deserno^{5,6,a} and Amy Orben^{1,a}

Abstract

Adolescence is a period of rapid change, with cognitive, mental wellbeing, environmental biological factors interacting to shape lifelong outcomes. Large, longitudinal phenotypically rich data sets available for reuse (secondary data) have revolutionized the way we study adolescence, allowing the field to examine these unfolding processes across hundreds or even thousands of individuals. Here, we outline the opportunities and challenges associated with such secondary data sets, provide an overview of particularly valuable resources available to the field, and recommend best practices to improve the rigor and transparency of analyses conducted on large, secondary data sets.

Addresses

¹ MRC Cognition and Brain Sciences Unit, University of Cambridge, UK

² Cognitive Neuroscience Department, Donders Institute for Brain, Cognition and Behavior, the Netherlands

³ Radboud University Medical Center, Nijmegen, the Netherlands

⁴ Department of Psychology, Institute for Psychiatry, Psychology and Neuroscience, King's College London, London, UK

⁵ Department of Psychological Methods, University of Amsterdam, Amsterdam, the Netherlands

⁶ Centre for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany

Corresponding author: Kievit, Rogier A (Rogier.Kievit@donders.ru.nl)

^a Contributed equally.

Current Opinion in Psychology 2022, 44:303–308

This review comes from a themed issue on **Adolescent Development**

Edited by **Lydia Krabbendam** and **Barbara Braams**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 14 October 2021

<https://doi.org/10.1016/j.copsyc.2021.10.003>

2352-250X/© 2021 Published by Elsevier Ltd.

Understanding adolescence using large secondary data sets

Adolescence is a period of profound and prolonged change in cognitive ability, mental wellbeing, environmental factors, social abilities, hormonal fluctuations,

and brain structure and function [10]. Large, deeply phenotyped, longitudinal data sets present a unique opportunity to understand the interplay between these multiple explanatory levels, capturing rich individual differences in processes that unfold during adolescence [24].

Large data sets offer numerous important benefits for understanding adolescent development. Such data sets expand our ability to ask research questions in ways often neglected by experimental studies, by improving our ability to study cross-cultural effects [13], diverse demographics [19], and historical and cohort differences [37]. They increase statistical power, safeguarding against misestimation of effect direction or magnitude (e.g. [20]), and allow us to examine differences in degree (e.g. factor models) or in kind (e.g. by identifying latent clusters or phenotypic profiles), [33] in outcomes of interest. Crucially, large longitudinal designs allow us to study antecedents of relatively rare but impactful phenotypes before their emergence, promising unique avenues toward identifying novel interventions. Although randomized controlled trials are often considered the gold standard for establishing intervention efficacy, they may not always be informative of developmental processes ‘in the wild’. For many developmental questions, carefully considered observational studies may in fact be the most sensitive and suitable tool to achieve such inferential aims [50]. Although such large, rich, data sets are often too costly in terms of time and logistics to collect at the individual or laboratory level, increasingly, they are made available as *secondary data sets* [26]: Data sets collected by other individuals, with other primary purposes in mind. Although secondary data sets may also be small (e.g. focused experimental data), and large data sets are not necessarily made available for secondary analysis, our focus here is on data sets that are both large (larger than typical in the field) and, in principle, publicly available.

Here, we outline different types of large secondary data sets, provide an overview of resources, and illustrate how we can test developmental theories and inform policy and practice, using the example of sensitive periods for adversity (Textbox 1). In addition, we highlight some

Textbox 1

Leveraging secondary data analysis to study sensitive periods for adversity

Sensitive periods refer to developmental phases of heightened sensitivity to environmental inputs (Knudsen, 2004). The study of sensitive windows is of foundational importance for research into adolescence, as it informs our understanding of the nature-nurture relationship at different developmental stages and guide the timing of interventions. However, the individual differences in timing and multivariate nature of sensitive windows places high demands on the data needed to improve our mechanistic understanding.

One of the best-studied areas of adversity is that of Adverse Childhood Experiences (ACEs), such as childhood abuse or neglect, which have been shown to have long-lasting consequences for later life outcomes (Felitti et al., 1998). These effects are particularly strong if they occur before the age of three (Dunn et al., 2019), providing strong evidence for early sensitive periods. The study of ACEs has fundamentally shaped the scientific and public understanding of childhood vulnerabilities, highlighting the potential impact of cohort studies: Since 2018, at least 37 US states have enacted or adopted legislation related to ACEs (NCSL, 2021). Although aspects of adversity can be informed through experimental work in rodents (e.g., Schneider & Koch, 2003), longitudinal effects of adversity on humans are mainly tractable through observational studies and natural experiments. Large cohorts with rich environmental data, such as ALSPAC, can be used to study how the timing and nature of a host of adversities relate to later life outcomes (Dunn et al., 2019; Andersen, 2021; Figure Textbox 1). To properly separate the correlated processes involved, large, rich developmental samples are needed. Other innovative approaches such as *regression discontinuity designs* (e.g., Hahn et al., 2001) allow us to leverage natural differences and changes in policies such as government intervention programs to isolate how environmental inputs impact individuals (Gong et al., 2020), or how additional education in late childhood and adolescence leads to lifelong increases in IQ (Ritchie & Tucker-Drob, 2018). Similarly, large twin design studies allow us to tease apart the causal interplay between genetic and environmental factors following adversity (e.g. Alemany et al., 2013, Shakoor et al., 2015).

We know much less about sensitive periods for adversity in adolescence (e.g., Fuhrmann et al., 2015), although emerging evidence suggests that different types of adversity, such as bullying or social exclusion, may be particularly harmful at this stage of development (Blakemore & Mills, 2014). Innovative new studies, such as the ABCD study, with deep, longitudinal phenotyping promise to help identify the processes which unfold prior to initial emergence of mental health and substance (ab)use problems. Eventually, these advances will help us to better quantify sensitive periods where (adverse) experiences have outsized impact on predicted life courses.

NCSL (2021). <https://www.ncsl.org/research/health/adverse-childhood-experiences-aces.asp>

1. Alemany, S., Goldberg, X., van Winkel, R., Gastó, C., Peralta, V., & Fañanás, L. (2013). Childhood adversity and psychosis: examining whether the association is due to genetic confounding using a monozygotic twin differences approach. *European Psychiatry*, 28(4), 207-212.
2. Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing?. *Annual review of psychology*, 65, 187-207.
3. Dunn, E. C., Soare, T. W., Zhu, Y., Simpkin, A. J., Suderman, M. J., Klengel, T., ... & Rellon, C. L. (2019). Sensitive periods for the effect of childhood adversity on DNA methylation: results from a prospective, longitudinal study. *Biological psychiatry*, 85(10), 838-849.
4. Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1), 201- 209.
5. Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study. *American journal of preventive medicine*, 14(4), 245-258.
6. Gong, J., Lu, Y., & Xie, H. (2020). The average and distributional effects of teenage adversity on long-term health. *Journal of health economics*, 71, 102288.
7. Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of cognitive neuroscience*, 16(8), 1412-1425
8. Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological science*, 29(8), 1358-1369.
9. Schneider, M., & Koch, M. (2003). Chronic pubertal, but not adult chronic cannabinoid treatment impairs sensorimotor gating, recognition memory, and the Orben, A., Lucas, R. E., Fuhrmann, D., & Kievit, R. (2020). Trajectories of adolescent life satisfaction. <https://psyarxiv.com/y8ruw> performance in a progressive ratio task in adult rats. *Neuropsychopharmacology*, 28(10), 1760-1769.
10. Shakoor, S., McGuire, P., Cardno, A. G., Freeman, D., Plomin, R., & Ronald, A. (2015). A shared genetic propensity underlies experiences of bullying victimization in late childhood and self-rated paranoid thinking in adolescence. *Schizophrenia bulletin*, 41(3), 754-763.

Leveraging secondary data analysis to study sensitive periods for adversity.

methodological challenges associated with large, secondary data sets, and provide an overview of best practices that can help improve the robustness and rigor of secondary data analysis.

Secondary data sets and where to find them

Funder mandates toward open data (e.g. the European Research Council (ERC) and the Wellcome Trust), the open science movement (e.g. [32]), and incentives such as 'open data' badges (e.g. [14]) have all encouraged increased data sharing and thus the availability of secondary data sets. These data sets differ in their scope and design but commonly include measures of demographics, mental health, mood, cognitive skills,

personality, and/or environmental factors relevant for understanding adolescent development. Increasingly, richer biological data are also collected, such as brain structure and function, genotyping, twin designs, hormone levels, and puberty status (e.g. [44]). The inclusion of these biological factors is especially impactful, as they often require high levels of expertise to acquire and process, as well as large sample sizes to allow their proper integration as explanatory factors, placing them out of reach for many smaller laboratories.

Cohort designs follow participants across time, for example, ALSPAC (Avon Longitudinal Study of Parents and Children), which tracks children born in

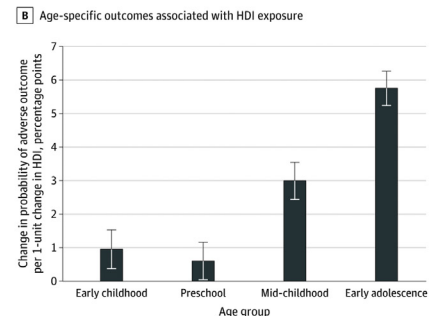


Figure Textbox 1. Changes in probability of adverse outcomes after household dysfunction (HDI) exposure at different ages in $N \sim 600,000$. The figure highlights that that sensitivity increases from early childhood to early adolescence. Figure adapted from Andersen (2021).

Figure 1

A	B	C	D	E	F	G	H
Datasets (with hyperlink)	Abbrev.	Type	Assessment/Topics	Waves	Spacing	Age	N (at baseline)
Adolescent Brain Cognitive Development Study	ABCD	birth/age cohort	Brain Imaging, Neurocognition, Mental & Physical health, Biospecimens	3	1 year	9-10 at baseline	11,868
Americans Changing Lives Study			Health outcomes, Psychosocial factors, Environment, Aging, Epigenetics	6	94/2001/2011/20	>25	3,617
Avon Longitudinal Study of Parents & Children	ALSPAC	birth/age cohort	Clinical measures, Biological Resources, Natural & Social Environmental	9 clinical a	irregular?	born 1991-1992	14,000
British Cohort Study 1970	BCS70	birth/age cohort	health, physical, educational and social development	9	1996/2000/2004	born in 1970	17,000
Centre for Attention Learning & Development	CALM	clinical population study	Cognition, Learning, Behavior, Neuroimaging, Genetics	2	3-5 years	5-18	863
China Family Panel Studies		household panel survey	economic activities, education outcomes, family dynamics and relationships	10	1 year	>9	30,000
Early Childhood Longitudinal Study	ECLS	kindergarten cohort	cognitive, social, emotional, and physical development, environment	9	1 year	5 at baseline	18,174
E-Risk Longitudinal Twin Study	E-Risk	birth/age cohort	mental health, obesity, asthma, school performance, criminal offending	5	age 5, age 7, age 10	age 5+, born in 1995	2,232
Fragile Families and Child Wellbeing Study		birth/age cohort	attitudes, relationships, parenting behavior, demographic characteristics	6	1, age 3, age 5, age 9	1 at baseline	4,897
The German Family Panel	pairfam	household panel survey	partnership, family development, child development, parenting	12	1 year	3 birth cohorts: 197	12,000

Current Opinion in Psychology

Screenshot of the Open Science Framework (OSF) resource on adolescent samples.

1991–1992, the Millennium Cohort Study, which tracks 19,000 United Kingdom children born in 2000–2002, or the ABCD (Adolescent Brain Cognitive Development) data set [44] which will track 11,000 United States 10-year-olds up to adulthood across one or multiple geographic locations, commonly at fixed intervals. In contrast, there also exist household panel studies, which are routinely operated at a national level and provide questionnaires to members of specific households over long periods of time (e.g. US (Understanding Society), YRBSS (Youth Risk Behavior Surveillance System), SOEP (German Socio-Economic Panel), HILDA (Household, Income and Labour Dynamics in Australia and PSID (Panel Study of Income Dynamics)). Increasingly, large-scale studies are collated into larger consortium data sets, for example, by combining diverse studies examining adolescents during the COVID-19 pandemic ([5]) or by harmonizing developmental cohorts that may include adolescent participants [46].

Data sets vary considerably in the richness and scope of their data, their ease of access (ranging from click-to-download to multimonth vetting processes), and their cost (from free-to-use to several thousand dollars). For the purpose of this article, we have created an Open Science Framework (OSF) resource (See Figure 1 for a screenshot) which collates a number of publicly available developmental data sets that include (parts of) adolescence, the domains and variables included, and details of their data access protocol. This document is dynamic and will be updated based on reader suggestions. The at-a-glance convenience of our spreadsheet is meant to complement, rather than substitute, existing, online platforms with extensive search functionality (e.g. <https://www.closer.ac.uk/>).

Methodological challenges in secondary data analysis

Above, we outline some of the substantive and methodological benefits of large secondary data sets in trying

to understand the mechanisms of development during adolescence. Although these data sets confer many benefits, there are also potential methodological challenges associated with their use. We briefly outline some of the most important and less appreciated challenges.

Large sample sizes mean that almost every statistical association will be statistically, if not practically, significant. A move away from reporting statistical significance and toward effect sizes and model comparison is therefore desirable [16]. In addition, the analytic flexibility offered by large, phenotypically rich data sets can lead to many different, yet defensible, analytic pipelines, [7,28,41] amplified even further if (raw) neuroimaging data are involved [9]. This challenge extends beyond individual studies to the fact that hundreds of researchers access the same data set over time, each with their own, often partially overlapping, research questions. Although Type I errors are often discussed within a single study (or the related idea of forking paths, [21]), multiple research teams investigating the same data set may also lead to inflated error rates [43].

Furthermore, large-scale studies must find a compromise between the competing goals of phenotypic breadth and practicable ease and length of administration. This means that measurements of a given domain, especially when not the core interest of a study, are likely to use short or simplified proxies, potentially compromising various forms of validity (e.g. [27,42]).

One of the biggest challenges in many open data sets is that study design idiosyncrasies may limit our developmental inferences. Most large cohort studies attempt to measure individuals (bi)annually with roughly equal spacing between measurements. This may be for reasons of logistical convenience, or a (mistaken, e.g. see [11]) assumption that longitudinal models cannot accommodate unequally spaced observations. As detailed in [31], cohort studies often confound the number of assessments (which may induce practice

effects), with the developmental effect of age and other developmental processes such as puberty (for details see [31]; Figure 7). However, researchers can use large sample sizes and/or tailored planned missingness sampling designs [51] to artificially decouple impacts of age, development, and/or practice. For instance, McCormick [31] described how subsampling based on pubertal development can be used to simultaneously model age and puberty.

A more metascientific question concerns the scientific desirability of many research studies relying on the same cohort. The United Kingdom has several popular data sets covering adolescence, including ALSPAC, US (Understanding Society), and MCS (Millennium Cohort Study). Although these data sets have many strengths, it is likely not desirable that most of our scientific understanding of adolescence relies to a great degree on a small subset of the population (let alone the global population, [19]). Furthermore, in studies with higher participant burden, notably Magnetic Resonance Imaging (MRI), external validity may be additionally reduced given the differences between MRI-compatible individuals and the general population [30].

Improving secondary data analysis

Although not all pitfalls can be addressed by individual researchers, many can be mitigated. We now offer some recommendations for best practices. First, we recommend researchers preregister their analyses whenever possible [45]. Although preregistration is more commonly performed before data collection, it is suitable for secondary analysis also, for instance, by having a timestamped data access track record [47]. Weston et al., [47] offer multiple recommendations for preregistration tailored to secondary data analysis and [49] offer a template to preregister multiple existing data sets into a single analysis within the domain of longitudinal studies of aging. An alternative to preregistering a single analytic plan is instead to run many or even all plausible analyses. This strategy, known as multiverse or specification curve analysis, allows a researcher to run, and report, the output of a wide variety of defensible analysis pipelines (e.g. in- or exclusion of covariates, parametric vs nonparametric analyses, and so on.) and report the full range of findings (e.g. [34]). Other forms of robustness checks include replicating the core analyses in an independent replication sample (which can be an independent subset of the full sample but ideally a fully independent data set), which allows one to compare the generalizability (or lack thereof) across similar or different adolescent cohorts (e.g. [17]).

Second, we recommend researchers report effect sizes as many associations may be statistically, but not practically, significant when analyzing large sample sizes. In doing so, we may need to adjust our collective expectations of what

Table 1

Secondary data analysis: Checklist.

	Yes/no
Is the study preregistered?	
Does it report robustness analyses/specification curve?	
Does it report effect sizes?	
Is there a replication sample (internal or external)?	
Does the article plot raw data?	
Does the article share analytic code?	

effect sizes to expect, and which ones to treat as substantial (e.g. [22]). Whether an effect size is ‘large enough’ is a context- and research-specific decision that needs to be considered and debated by the community [4].

Relatedly, we recommend that researchers always plot the raw data in their articles. Summary plots such as bar graphs, summary statistics such as the mean/SD (Standard Deviation), or estimates such as correlation coefficients in isolation obscure the richness of distributions, the presence of floor or ceiling effects, the likelihood of subgroup heterogeneity, the appropriateness of parametric techniques, and many more features offered by inspection of raw data [48]. Although visualizing raw high dimensional data may be challenging, tailored approaches such as raincloud plots [2] and related techniques [8] can easily overcome practical challenges.

Finally, the analyses conducted on secondary data are commonly more complex than those applied to simpler experimental designs [28]. Methods sections in high-impact journals are often highly condensed or hidden at the end of an article, which can make it difficult or even impossible to assess which analyses exactly were performed. To address this issue, we recommend authors always publish the full analytic code, even when the raw data cannot be directly shared (e.g. [27,35]). This allows reviewers to assess precisely which analyses were completed, replicate the findings by applying for the data, and ideally expand on the work in future analyses.

Discussion

Secondary data offer a powerful resource to study the mechanisms and outcomes associated with (individual differences in) adolescent development. These samples increase power, allow for a considerably wider range of analytic strategies, and can allow us to ask questions not accessible using classic, smaller-scale experimental designs. Crucially, they improve access to uniquely rich data sets across laboratories and research environments that differ in resources. We provide an easily accessible overview of secondary data sets

researchers may want to explore. In addition, we outline some of the unique challenges that can come along with large secondary data sets and how they can be ameliorated to a considerable degree by increased awareness and better practices (see Table 1) that help improve robustness and transparency. Other meta-scientific challenges of secondary data sets remain to be resolved by the field, including the consequences of over-reliance on a small number of data sets, and the ethical challenges associated with (degrees of) anonymity offered by increasingly rich data sets. By tackling these challenges transparently, secondary data sets offer a unique, complementary source of scientific inquiry into the complex, but fascinating, processes that unfold during adolescence.

Funding

AO was supported by a College Research Fellowship from Emmanuel College, University of Cambridge and UK Economic and Social Research Council ES/T008709/1. MKD was supported by a Rubicon fellowship from the Netherlands Organization for Scientific Research (NWO) no.019.191SG.005. DF was supported by the UK Economic and Social Research Council ES/T015861/1. RAK was supported by a Hypatia Fellowship at the RadboudUMC.

Conflict of interest statement

No conflict of interest.

Acknowledgements

The authors like to thank Georg Henning and Laura Havers for valuable input on our previous draft.

References

Papers of particular interest, published within the period of review, have been highlighted as:

- * of special interest
- ** of outstanding interest

2. Allen M, Poggiali D, Whitaker K, Marshall TR, Kievit RA: **Rain-cloud plots: a multi-platform tool for robust data visualization.** *Wellcome Open Research* 2019, **4**.
5. Barendse M, Flannery J, Cavanagh C, Aristizabal M, Becker SP, Berger E, Pfeifer J: **Longitudinal change in adolescent depression and anxiety symptoms from before to during the COVID-19 pandemic: an international collaborative of 12 samples.** 2021. <https://psyarxiv.com/hn7us/>.
7. Botvinik-Nezer R, Holzmeister F, Camerer CF, Dreber A, Huber J, Johannesson M, Rieck JR: **Variability in the analysis of a single neuroimaging dataset by many teams.** *Nature* 2020, **582**:84–88.
8. Bulut & Desjardins: **Exploring, visualizing, and modeling big data with R.** 2019. <https://okanbulut.github.io/bigdata/>.
9. Carp J: **On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments.** *Front Neurosci.* 2012, **6**:149.
10. Dahl RE, Allen NB, Wilbrecht L, Suleiman AB: **Importance of investing in adolescence from a developmental science perspective.** *Nature* 2018, **554**:441–450.
11. de Haan-Rietdijk S, Voelkle MC, Keijsers L, Hamaker EL: **Discrete-vs. continuous-time modeling of unequally spaced experience sampling method data.** *Front. Psychol.* 2017, **8**:1849.
13. Duell N, Steinberg L, Icenogle G, Chein J, Chaudhary N, Di Giunta L, ... Chang L: **Age patterns in risk taking across the world.** *J. Youth Adolesc.* 2018, **47**:1052–1072.
14. Eich Eric: **Business not as usual.** *Psychol Sci* 2014:3–6.
16. Fritz CO, Morris PE, Richler JJ: **Effect size estimates: current use, calculations, and interpretation.** *J. Exp. Psychol.* 2012, **141**:2–18, <https://doi.org/10.1037/a0024338>.
A primer on effect size use and interpretation
17. Fuhrmann D, Simpson-Kent IL, Bathelt J, , CALM Team, Kievit RA: **A hierarchical watershed model of fluid intelligence in childhood and adolescence.** *Cerebr. Cortex* 2020, **30**:339–352.
19. Ghai S: **It's time to reimagine sample diversity and retire the WEIRD dichotomy.** *Nature Human Behaviour* 2021:1–2.
20. Gelman A, Carlin J: **Beyond power calculations: assessing type S (sign) and type M (magnitude) errors.** *Perspect. Psychol. Sci.* 2014, **9**:641–651.
21. Gelman A, Loken E: **The garden of forking paths: why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time.** Department of Statistics, Columbia University; 2013:348.
22. Gignac GE, Szodorai ET: **Effect size guidelines for individual differences researchers.** *Pers. Individ. Differ.* 2016, **102**:74–78.
24. Guez A, Peyre H, Williams C, Labouret G, Ramus F: **The epidemiology of cognitive development.** *Cognition* 2021:104690.
26. Johnston MP: **Secondary data analysis: a method of which the time has come.** *Qualitative and Quantitative Methods in Libraries* 2017, **3**:619–626.
27. Kievit RA, Fuhrmann D, Borgeest GS, Simpson-Kent IL, Henson RN: **The neural determinants of age-related changes in fluid intelligence: a pre-registered, longitudinal analysis in UK Biobank.** *Wellcome Open Research* 2018, **3**.
28. King KM, Littlefield AK, McCabe CJ, Mills KL, Flournoy J, Chassin L: **Longitudinal modeling in developmental neuroimaging research: common challenges, and solutions from developmental psychology.** *Developmental Cognit. Neurosci.* 2018, **33**:54–72.
A useful overview of longitudinal analysis strategies.
30. LeWinn KZ, Sheridan MA, Keyes KM, Hamilton A, McLaughlin KA: **Sample composition alters associations between age and brain structure.** *Nat. Commun.* 2017, **8**:1–14.
31. McCormick EM: **Multi-Level Multi-Growth Models: new opportunities for addressing developmental theory using advanced longitudinal designs with planned missingness.** *Developmental Cognit. Neurosci.* 2021, **51**:101001.
Paper demonstrating how standard study designs conflate correlated processes of biological maturity (e.g. puberty), calendar age and practice/retest expertise effects. Offers solutions that leverage interval heterogeneity
32. Munafò Marcus R, Nosek Brian A, Bishop Dorothy VM, Button Katherine S, Chambers Christopher D: **Nathalie percie du sert, uri simonsohn, eric-jan wagenmakers.** In *A manifesto for reproducible science.* Edited by Ware Jennifer J, Ioannidis John PA, **1**; 2017:1–9. *Nature Human Behaviour*.
33. Oberski D: **Mixture models: latent profile and latent class analysis.** In *Modern statistical methods for HCI.* Cham: Springer; 2016:275–287.
34. Orben A, Przybylski AK: **The association between adolescent well-being and digital technology use.** *Nature Human Behaviour* 2019, **3**:173–182.

Paper using a multiverse/specification curve approach to examine the robustness of associations to multiple defensible analytic pipelines.

35. Orben A, Lucas RE, Fuhrmann D, Kievit R: *Trajectories of adolescent life satisfaction*. 2020. <https://psyarxiv.com/y8ruw>.
 37. Platt JM, Keyes KM, McLaughlin KA, Kaufman AS: **The Flynn effect for fluid IQ may not generalize to all ages or ability levels: a population-based study of 10,000 US adolescents**. *Intelligence* 2019, **77**:101385.
 41. Simmons JP, Nelson LD, Simonsohn U: **False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant**. *Psychol. Sci.* 2011, **22**:1359–1366.
- Influential paper on the impact of analysis flexibility.
42. Taylor BK, Frenzel MR, Eastman JA, Wiesman AI, Wang YP, Calhoun VD, ... Wilson TW: **Reliability of the NIH toolbox cognitive battery in children and adolescents: a 3-year longitudinal examination**. *Psychol. Med.* 2020:1–10.
 43. Thompson WH, Wright J, Bissett PG, Poldrack RA: **Meta-research: dataset decay and the problem of sequential analyses on open datasets**. *Elife* 2020, **9**, e53498.
 44. Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, ... Weiss SR: **The conception of the ABCD study: from substance use to a broad NIH collaboration**. *Developmental Cognit. Neurosci.* 2018, **32**:4–7.
- One of the largest available studies on developmental cognitive neuroscience.
45. Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HL, Kievit RA: **An agenda for purely confirmatory research**. *Perspect. Psychol. Sci.* 2012, **7**:632–638.
 46. Walhovd KB, Fjell AM, Westerhausen R, Nyberg L, Ebmeier KP, Lindenberger U ... , Lifebrain Consortium: **Healthy minds 0–100 years: optimising the use of European brain imaging cohorts (“Lifebrain”)**. *Eur. Psychiatr.* 2018, **50**:47–56.
 47. Weston SJ, Ritchie SJ, Rohrer JM, Przybylski AK: **Recommendations for increasing the transparency of analysis of preexisting data sets**. *Adv. Methods and Pract. in Psychol. Sci.* 2019, **2**:214–227.
- An overview of best practices for robust science in secondary datasets.
48. Weissgerber TL, Milic NM, Winham SJ, Garovic VD: **Beyond bar and line graphs: time for a new data presentation paradigm**. *PLoS Biol* 2015, **13**, e1002128.
 49. Willroth EC, Graham EK, Mroczek DK: **Challenges and opportunities in preregistration of coordinated data analysis: a tutorial and template**. *Psychol. Aging* 2021.
 50. Zyphur MJ, Allison PD, Tay L, Voelkle MC, Preacher KJ, Zhang Z, ... Diener E: **From data to causes I: building a general cross-lagged panel model (GCLM)**. *Organ. Res. Methods* 2020, **23**:651–687.
 51. Little TD, Rhemtulla M: **Planned missing data designs for developmental researchers**. *Child Dev Perspect* 2013, **7**: 199–204. Chicago.