HHS Public Access

Coordinated Data Analysis: Knowledge Accumulation in Lifespan Developmental Psychology

Eileen K. Graham[1], **Emily C. Willroth**[1], **Sara J. Weston**[2], **Graciela Muniz-Terrera**[3,4], **Sean A.P. Clouston**[5], **Scott M. Hofer**[6], **Daniel K. Mroczek**[1,7], **Andrea M. Piccinin**[6]

[1]Department of Medical Social Sciences, Northwestern University

[2]Department of Psychology, University of Oregon

[3]University of Edinburgh, Centre for Clinical Brain Sciences, Edinburgh, Scotland, UK

[4]University of Edinburgh, Centre for Dementia Prevention, Edinburgh, Scotland, UK

[5]Department of Family, Population, and Preventive Medicine; Program in Public Health, University of Stony Brook

[6]University of Victoria, Department of Psychology, Victoria, BC, Canada

[7]Department of Psychology, Northwestern University

Abstract

Coordinated analysis is a powerful form of integrative analysis, and is well suited in its capacity to promote cumulative scientific knowledge, particularly in subfields of psychology that focus on the processes of lifespan development and aging. Coordinated analysis uses raw data from individual studies to create similar hypothesis tests for a given research question across multiple datasets, thereby making it less vulnerable to common criticisms of meta-analysis such as file drawer effects or publication bias. Coordinated analysis can sometimes use random effects meta-analysis to summarize results, which does not assume a single true effect size for a given statistical test. By fitting parallel models in separate datasets, coordinated analysis preserves the heterogeneity among studies, and provides a window into the generalizability and external validity of a set of results. The current paper achieves three goals: First, it describes the phases of a coordinated analysis so that interested researchers can more easily adopt these methods in their labs. Second, it discusses the importance of coordinated analysis within the context of the credibility revolution in psychology. Third, it encourages the use of existing data networks and repositories for conducting coordinated analysis, in order to enhance accessibility and inclusivity. Subfields of research that require time- or resource-intensive data collection, such as longitudinal aging research, would benefit by adopting these methods.

Correspondence concerning this article should be addressed to Eileen K. Graham, 625 N. Michigan Ave, Chicago IL 60611, eileen.graham@northwestern.edu.
All authors contributed to the conceptualization, preparation, and revision of this manuscript.

Introduction

Concerns about replicability, credibility, and cumulative science have impacted nearly all subfields of the psychological sciences and other related sciences (Nelson, Simmons, & Simonsohn, 2018; Rodgers & Shrout, 2018). It is clear, however, that the varying needs of different research areas make it impossible to apply a "one size fits all" approach to these issues. Lifespan development, health psychology, and aging are cases in point: replication is a core scientific value, yet, replications are difficult to accomplish where complex longitudinal or time-intensive data are needed (Hofer & Piccinin, 2009, 2010; Piccinin & Hofer, 2008; Weston, Graham, & Piccinin, 2019) or where participants are followed for long periods until a certain event occurs such as onset of a disease, a relationship event (like divorce), or death. Thus, fields that rely heavily on long-term longitudinal data, such as aging research, lifespan development, some subareas of health psychology, behavioral and social epidemiology, and others, face unique challenges when trying to adopt open, reproducible, and replicable research practices.

This paper discusses coordinated analysis as one approach to addressing some of these issues. Coordinated analysis (Hofer & Piccinin, 2009) can be a highly useful approach for replication, synthesis, and evaluation of heterogeneity of results from complex longitudinal research, as has been demonstrated in the Integrative Analysis of Longitudinal Studies of Aging and Dementia (IALSA) research network. Coordinated analysis is a methodological approach in which optimally similar statistical models are estimated on datasets independently with code designed to maximize measurement similarity across samples, and there is an individual set of results for each dataset with a possible accompanying meta-analysis. "Integrative," in the context of the IALSA network and approach, refers to three key goals: 1) multivariate integration across domains of study (e.g., health, cognition, personality; Hofer & Piccinin, 2010), 2) integration of results across independent studies that differ in country, cohort, variables, and design (Hofer & Piccinin, 2009; Piccinin & Hofer, 2008), and 3) integration across levels of analysis (i.e., across research designs and statistical methods: between-person vs. within-person results, patterns of change, mean trends; Hofer & Sliwinski, 2006; Hofer & Piccinin, 2010). Below, we describe how coordinated analysis, which can be implemented in service of each of these goals, is a powerful way to meet the challenges of replication in longitudinal research.

The current paper has three objectives: First, to provide a *description* of the usual phases of a successful coordinated analysis research project and make specific *prescriptions* for optimal approaches, including a worked example of a project currently underway. Second, to discuss the history and importance of coordinated analysis to the accumulation of knowledge in lifespan development. Third, to introduce how the use of existing data networks and

repositories to conduct coordinated analysis can be used by all researchers to accelerate the accumulation of knowledge.

Phases of a Coordinated IDA

The steps to completion of a coordinated analysis closely mirror the phases of a registered report (Chambers, Dienes, McIntosh, Rotshtein, & Willmes, 2015), in which the investigator pre-plans and pre-registers their introduction, approach, method, materials, and analysis plan (Weston et al., 2019; Willroth et al., in press). This workflow is optimal for any research endeavor, and the following sections describe how it applies in the context of coordinated analysis. We also emphasize the importance of pre-registering and time-stamping the specific questions, hypotheses, and detailed analytic plan prior to data access but note that in working with multiple teams using a distributed approach, this process becomes a part of the workflow. Given that many of these specifics follow an idealized scientific method, the basic process of a coordinated analysis will look familiar.

Each of the following sections contains a description of the essential steps within a given phase, followed by a discussion of challenges and considerations that one may encounter during that phase, along with our recommendations for possible solutions. Lastly, we will illustrate our description of each phase with a worked example of a coordinated analysis project currently underway in our research group, as of this writing (Willroth et al., in prep). The pre-registration, R scripts, and instructions sent to analysts for this coordinated analysis can be found on the Open Science Framework at https://osf.io/kwcd7/. See also Figure 1 for a flow chart outlining the basic phases and products yielded from each phase.

Phase 1: Define the Research Question and Conduct a Literature Search

The first phase in a coordinated analysis, as in other research endeavors, is to define the research question and conduct a thorough literature search. As the research questions are fine-tuned and the literature search is completed, be sure to include papers published using the specific datasets that are under consideration for use in the current project. In addition, findings from datasets that are not available for the proposed project are important to review, to ensure as comprehensive a summary as possible. The investigator should become familiar with key publications, including datasets, models, constructs, covariates, and data transformations used. The goal is to define simple and foundational research questions to be answered, while also preparing to replicate other aspects of the questions addressed by the established literature. When approached in this way, a precise set of research questions will ultimately be posed, with corresponding hypotheses and predictions. This process of searching the literature and developing research questions/hypotheses will be iterative, as the researcher develops a deeper knowledge of past findings and fine tunes the details of the project questions.

A key consideration during this phase is the complexity of the research question being posed. As we suggest in Weston, Graham, and Piccinin (2019) it is often best to choose a simple yet foundational question. With complex questions and models that are increasingly common in aging research, many datasets are unlikely to fulfill highly specialized variable, design, or data requirements required to test them (Kelly et al., 2016).

A single dataset project may involve quite complicated analyses that are fairly common in the aging literature, perhaps testing multiple mediational pathways, or longitudinal models testing several alternative time metrics (e.g., time-in-study, time-to-death, age). As analysis complexities increase, however, so do concerns regarding both replicability and generalizability. Models with many parameter estimates or complex structure will invariably encounter convergence issues when tested on a variety of datasets. For example, an auto-regressive latent trajectory model has very heavy and specific data demands and is typically quite arduous to compute even in a single dataset. We urge caution when posing questions requiring highly complex models, and suggest focusing on simpler models. A trade-off between answering a simple question with many datasets and answering a complex question with only a few may be necessary, with the balance of the decision depending on the specific topic. However, if the data specifications are so complex that most datasets do not meet the requirements, then it can become more difficult to draw conclusions from a coordinated analysis. For example, if only two studies with comparable study designs are included, and the analyses yield conflicting results, then interpreting these results and drawing a confident conclusion becomes difficult.

In short, using the most parsimonious model possible to answer your question is particularly important for coordinated analysis for two key reasons, one practical and the other philosophical. Practically speaking, it is much easier to harmonize, estimate, interpret, and compare results from a simpler model than a more complex model. Philosophically speaking, the credibility revolution is relying on researchers to confirm the reliability of the broader findings in our field (Vazire, 2018). In order to effectively do this, the first generation of coordinated analyses in aging research would do well to first address the simple associations and patterns.

Another challenge that may arise is the discovery that one or more datasets being considered for analysis has already been used to address a similar question. Some may consider dropping these datasets from the coordinated analysis, and include the results from those studies in the introduction and study justification only. On the contrary, we contend that the inclusion of these datasets is important for the following reasons: First, the reanalysis of those datasets provides information about the reproducibility of the findings (Condon, Graham, & Mroczek, 2017), and addresses whether the effects found in the original paper are robust to differences in modeling approach, covariates, or data transformations. Second, if the data are part of an ongoing data collection project (e.g., large-scale longitudinal studies typical in aging research), it is likely that more data have been added since the original paper was published. This provides researchers the opportunity to conduct a within study replication. That is, within the same study but with additional data (e.g., additional mortality data, additional disease event occurrences, new measurement occasions of a key variable), do the previously reported findings replicate when these new data are analyzed? Third, by including these studies in the coordinated analysis with the harmonized modeling approach, a more precise meta-analysis, if implemented, can be conducted. Relatedly, excluding those datasets is in itself a form of systematic exclusion, since non-null findings are more likely to have been published. By including these datasets in the coordinated analysis and harmonizing the analytic decisions, these questions are examined and articulated,

and comparisons across studies (or results from a meta-analysis) can be interpreted more confidently.

Example Application of Phase 1: Define the Research Question and Conduct a Literature Search—In an ongoing coordinated analysis that we are using here as illustration, we defined our research question: How is change in personality associated with mortality risk? This question follows two coordinated analyses previously conducted by our research group: Graham et al., (2017) examined the effects of personality levels on mortality risk and Graham et al., (2020a) examined average change in personality across the adult lifespan. We compiled an initial list of samples that met the data requirements for addressing this question, that would be further fine tuned in phase 3 (below). The initial iteration of this list included 13 samples. This list was compiled based on searching the data dictionaries of studies affiliated with the Integrative Analysis of Longitudinal Studies of Aging and Dementia (IALSA) research network and by cross-referencing prior coordinated analyses conducted by our research group (Graham et al., 2017, 2020a).

After defining the research question and identifying an initial list of samples, we conducted a literature search. This search yielded papers using some of the specific samples that we planned to include in our coordinated analysis (Mroczek & Spiro, 2007; Sharp, Beam, Reynolds, & Gatz, 2019). This is one example of why it is important to conduct a thorough literature search prior to pre-registering and conducting the coordinated analysis. In our search, we did not identify any papers that directly addressed our specific research question using samples outside of those we planned to include. However, if we had found such papers, we would have investigated whether we could add those samples to our coordinated analysis.

Phase 2: Develop an Analysis Plan

The goal of the second phase is to develop an analysis plan to appropriately answer the research questions and test all hypotheses. Given that prior papers will have estimated models that vary a great deal from one another (even if intending to answer the same question), this phase is critical. Investigators have the opportunity to build upon prior work and develop an optimal set of models that will test their hypotheses and answer their questions. For any analysis, the details and decision points will be unique to the specific research questions and content area, but the key to a successful coordinated analysis is the harmonization of models across all datasets. This means that each dataset uses (to the extent possible) identical code to fit all models. Using scripts that can be automated and deployed in a parallel manner across many datasets will help streamline this process. During this phase, researchers will need to specify every detail of the analysis, including cleaning, participant-level exclusions, variable transforming, restructuring, and modeling. The modeling plan should include the models that will test each hypothesis, as well as sensitivity analyses and robustness checks, such as adjusted and unadjusted models, accounts for attrition, a plan for handling outliers and other supplemental analyses that are relevant to aging research. By pre-registering as much of both the primary and sensitivity analyses as possible in advance (acknowledging that some methods [e.g., factor and mixture models] require iterative, data-dependent, decisions) and transparently reporting

any decisions that were made after pre-registration, researchers increase confidence in the credibility of their work by demonstrating that any supplemental analyses were not planned after the results were known (Weston et al., 2019; Willroth et al., in press). The analysis plan should also include blueprints for summarizing the individual study results, including visualizations and possible meta-analyses.

In addition to a narrative comparison of individual study results (e.g., Bendayan et al., 2017; Brown et al., 2012, Cadar et al., 2016; Lindwall et al., 2012; Mitchell et al., 2012; Scott, et al., 2020; Stawski et al., 2019; Yoneda et al., 2020a), meta-analysis can be easily applied to the planned results synthesis, and can be done so appropriately as a part of the planned comparisons of the individual study results when the studies are sufficiently comparable. Average effects can be estimated using a prescribed weighting convention, as well as heterogeneity statistics (e.g. I^2 and Cochran's Q (Patsopoulos, Evangelou, & Ioannidis, 2008)). The random effects meta-analytic model does not assume that one true effect size underlies all available studies, assuming instead that there may be many different true effect sizes corresponding to different populations (Borenstein, Hedges, Higgins, & Rothstein, 2010; Hedges & Vevea, 1998). Thus, any observed heterogeneity at this stage of the coordinated analysis should be viewed in light of this assumption for the group of studies being analyzed. Study level factors identified during the analysis planning and pre-registration phase may account for some of the between-study variation in effects. Some examples of between-study differences that may give rise to variations in effect sizes and are relevant to psychological aging questions are: country of origin, scale or measures used, language of measure, sample age/cohort (or age at entry into study), baseline year, and variation in spacing of measurement occasions. Additionally, these factors themselves may vary in importance depending on the nature of the research questions being addressed. The scripts to summarize results across studies can include both basic meta-analysis as well as meta-regression models that include these moderators. The meta-analytic plan should be pre-registered along with the individual study analytic plan, and should specify which estimates from each model will be meta-analyzed (e.g. the linear slope estimate, specific interaction terms), which heterogeneity statistics will be interpreted, and which study-level variables will be entered into these models to account for heterogeneity.

When creating an analysis plan for a coordinated analysis, three main levels of harmonization need careful consideration: variable harmonization, model harmonization, and study-design harmonization. *Variable harmonization* begins with identifying operational definitions of key constructs. This includes identifying the specification, within each dataset, of the predictor and outcome variables, as well as moderators (or mediators), time metrics, and possible covariates. When selecting constructs for the analysis, specifying the degree of flexibility allowable in these operational definitions is essential. Researchers will inevitably encounter measurement differences in constructs across studies (e.g., Bendayan, et al., 2020; Brown et al., 2012; Duggan et al., 2019; Griffith et al., 2015; Griffith et al., 2013; Hoogendijk et al., 2020; Lindwall et al., 2012; Mitchell et al., 2012; Yoneda et al., 2020b; Zammit et al., 2019), even when the instrument used was the same across studies (Piccinin et al., 2013a). Measurement harmonization is a common challenge when coordinating across multiple longitudinal studies, because measures are often modified in order to reduce participant burden. For example, several studies may appear to have

used the same instrument to measure a construct, but some studies used an ordinal or continuous scale, while others used a binary indicator. One solution would be to transform the continuous scale into a simple binary indicator, but this is clearly not optimal as it masks the rich variability from the continuous measure. If only a small number of datasets use the binary measure, one option is to test multiple models across a subset of your studies and compare them. Considering existing harmonization tools, such as the ISCED for education classification, is recommended, though these may not always be helpful (Piccinin et al., 2013a). Conceptual variable harmonization challenges are not unsurmountable, but need to be resolved in order to ensure optimal comparability. Many of these decisions will be made on a case by case basis, and the most important piece is to make these decisions prior to analyzing any of the datasets.

A useful tool to create while harmonizing variables is a detailed combined data dictionary for the specific variables from each study that will be used. This data dictionary can include exact variable names and variable types (e.g. ordinal, continuous, count, binary), response options, and variable labels for each study, as well as a clear harmonization/transformation plan to ensure that each construct is coded as comparably as possible (Fortier et al., 2016). Once these decisions have been made, it is time to write the code, script out the desired data structure (e.g. long vs. wide form), determine variable coding and transformation, which items should be included in composites, determine exact model specifications, and develop a clear and executable plan for summarizing, plotting/visualizing, tabling, and synthesizing the results.

Model harmonization is essential in coordinated analysis because it gives researchers greater confidence that differences in individual study results are not due to differences in parameterization of the models themselves. The best way to ensure that the models are harmonized across studies is to create a script that streamlines all analytic decisions and can be used across datasets with minimal input from the data analysts. Model harmonization can be particularly challenging within the context of aging research, especially when the research topic requires a complex longitudinal data structure such as a measurement burst design. Certain aging research questions involve the disentangling of age-, history-, and cohort-effects, and require complex models (e.g., bi-directional or auto-regressive models, testing of competing models with alternative time metrics). As discussed previously, a coordinated analysis of complex designs and models will usually lead to a smaller number of available datasets, but can ultimately be the necessary solution when the only way to appropriately test the corresponding hypotheses is with these more complex models.

Study design harmonization is the least plausible of these three types of harmonization in the context of secondary analysis, but considering differences in study designs is of equal importance for understanding and accounting for variations in results. While it is possible to develop a harmonized modeling plan, it is impractical to expect true harmonization at the study design level. Most questions about aging processes inquire about the impact of within-person status and change in various factors (e.g. health, cognition), and all studies vary on design features such as measurement spacing, developmental period of the sample, or simply the year that the study was initiated, and these may influence the construction and interpretation of the models. That said, it might sometimes be possible for specific

analytic and model selection decisions to at least partially account for these cross-study variations. For example, if a coordinated analysis project is using continuous time modeling for panel data, one could plan a comparison of parameter estimates for models with different measurement lags. This allows the inherent design differences to remain but allows for comparison of these differences (e.g., the impact of retest effects at 1 vs. 2+ year intervals). Whether this occurs at the individual study analysis level, or at the meta-analytic level, is ultimately up to the investigative team. In a coordinated analysis, we recommend coding these study level differences and considering them as moderators in a narrative comparison or meta-analytic phase (meta-regression), to evaluate the extent to which study-level differences might account for variations in the estimates (or other features of the model). See also Piccinin et al., (2011) for a detailed discussion of the myriad technical considerations of the various analytic approaches to the study of age-related change. These considerations should be tailored to the specific research questions, and all of them can be used within the context of coordinated analysis. The variety of design characteristics in each study should be leveraged in order to examine sources of heterogeneity in the results. When the models themselves are harmonized to the extent possible, the results will be more directly comparable, and differences in study-level design features (e.g., number and spacing of measurement occasions, age range of the sample, baseline year) can be included at the meta-analytic level. In this way, variations in study design can be a potential asset rather than a liability.

Example of Application of Phase 2: Develop an Analysis Plan—After refining our research question (are changes in personality traits associated with mortality?), conducting a thorough literature search, and identifying a preliminary list of datasets, we moved to phase 2. Our analysis plan was informed in part by the methods used in the single-sample papers identified in our literature search in Phase 1. We amended those methods to address current best-practices and to flexibly accommodate the characteristics of our samples. The first step of our analysis plan was to operationalize our key predictor (personality change) and our key outcome (mortality), and to develop a variable harmonization plan. We operationalized personality as Big Five personality traits assessed using any multi-item published Big Five measure. We could have chosen a narrower operationalization of personality traits, but this would have dramatically reduced the number of samples that we could include. Likewise, we could have chosen a broader operationalization of personality (e.g., any personality facet or trait-like measure, such as hostility); however, this approach would have resulted in a large, unwieldy analysis with high levels of heterogeneity. We also specified the minimum number of measurement occasions required to reliably compute longitudinal change in personality. We operationalized mortality as time from baseline personality measurement to death, with death status and month and year of death obtained from any reliable source such as a national mortality database. Given the inherent complexity of our research question and the large number of included samples, we decided not to include mediators or moderators.

Next, we drafted a pre-registration using the OSF template for secondary data analysis (Weston et al., 2019; Willroth et al., in press). Our pre-registration included the research question and hypotheses, operationalizations of study variables (both the general rules described above and the specific measures used in each sample), details about variable

transformations, handling of outliers and missing data, and specification of the exact statistical models, in order to achieve optimal variable and model harmonization. Notably, the analysis plan included a plan for summarizing meta-analytic results. We decided that meta-anlaysis would be a useful way to interpret the results for this project, so we planned to use random-effects meta-analysis to calculate the overall weighted mean effect sizes, standard errors, and 95% CIs around effects. To address study-level harmonization, we also detailed a plan for examining heterogeneity in effects and testing three potential study-level moderators that might explain that heterogeneity (average baseline age, country-of-origin, measurement differences). Although we drafted the complete pre-registration in Phase 2, we did not finalize or publicize it until the end of Phase 3, once the list of studies was finalized.

After drafting our analysis plan, we created three R scripts to carry out the planned analyses. The first script contained code to clean and prepare the datasets. This script included only the cleaning steps that could not be fully automated. The second script contained code to carry out the individual study analyses and return an output object. This analysis script was fully automatized such that it could be applied to any dataset that had been prepared with the data preparation script, without editing the code. We programmed this script to store several pieces of information into an output object: (1) characteristics of the sample such as baseline age, country, sample size, operationalization of personality, operationalization of mortality; (2) descriptive statistics of the study variables and participant demographics; and (3) results from all models, including simpler unadjusted versions of our final covariate-adjusted models. The third and final script contained code to carry out random effects meta-analyses, visualize hazards plots, forest plots, and create tables. The written analysis plan (in the form of the pre-registration) and the coded analysis plan (in the form of the three R scripts) were iterative as making changes to one necessitated making changes to the other.

Phase 3: Final Study Selection

Third, the list of datasets that contain the requisite data to fit the proposed models is fine tuned, and contact is initiated with study investigators. Along with documenting study characteristics, it is also necessary to evaluate the limitations and strengths of each dataset in terms of its ability to address the research question. Many datasets will be publicly available, some will require data use agreements and an application process, while others will not be directly accessible and will require coordination with a study-site analyst or server for remote analysis (e.g., DPUK data portal). During this phase you will confirm which of these categories each dataset falls into, establish a means of communication with the PI's and analysts, and reach an agreement with study sites about authorship expectations.

As researchers proceed through phase 3, it is important to revisit the research question to continually evaluate feasibility and whether additional questions are warranted. In this way, the first three phases of a coordinated analysis will be *iterative* as the questions, analysis plan, and selected studies are fine-tuned. Some of the issues that were addressed in Phase 2 will need to be revisited in Phase 3. For example, as part of the check on dataset and research question compatibility, investigators will need to re-evaluate the potential for variable, model, and study level harmonization. Differences in variable operationalization cannot easily be separated from lack of replicability. Similarly, study design features such as

follow-up time along with number and spacing of waves (e.g., 4 biannual waves versus 10–15 yearly waves), and sample characteristics such as size, developmental period, age range, and demographic heterogeneity are essential details to document during study selection. These longitudinal design differences can be particularly impactful in psychological aging research, and could be accounted for either within each study (e.g., as dummy codes or covariates) during the individual study analysis, or as moderators at the meta-analytic stage.

Another important consideration during this phase is statistical power. Publications that have used coordinated analysis can have very high total sample sizes, but some of the individual studies can still be relatively small. While many datasets in aging research have samples of well over 1,000 participants, power can still be an issue depending on the specific question. For example, if the analytic plan requires multiple measurement occasions of a single construct both before and after an event, the number of participants within a given sample that have both experienced the event *and* have multiple measurements after the event can be quite low, especially if the event in question lends itself to high attrition. This can lead to inflated estimates and unreliable models, and should be interpreted cautiously. Often, this level of information is unknowable until the analytic phase, once the data have been accessed, so the contingency plans should be strategized during phase two. Ultimately, the study-level inclusion criteria are at the discretion of the project leader, but the study-level and participant-level inclusion criteria should be planned in advance. Any deviations from planned inclusion criteria that are guided by actual results (once known) should be transparently reported.

As Phase 3 is completed, and prior to data access, pre-registration of the project is *strongly* recommended in order to enhance the credibility of the process (Van den Akker et al., 2019), using the template for secondary data analysis (Weston et al., 2019; Willroth et al., in press). The more studies included in a coordinated analysis, the more time it will ultimately take to plan and analyze, but also the more potentially robust set of findings in the end. In sum, this approach to longitudinal aging research will seem slower in the short term, but can ultimately be more impactful. A set of 10 to 15 findings with meta-analytic summary (see (Wood et al., 2018) for an example using 83 studies) will transcend the typical one or two study manuscript in both authoritativeness and credibility (Vazire, 2018).

Application of Phase 3: Final Study Selection—We repeated our search for relevant samples to ensure we had not missed any, and initiated data requests for those that were not publicly available. Through this process, we were notified of two additional samples that fit inclusion criteria, highlighting the need to wait until the end of Phase 3 to finalize and publicize the pre-registration. After adding these new samples to our documentation and receiving permission to use the other samples, we posted the pre-registration and R scripts on the Open Science Framework.

Phase 4: Analysis and Model Summarizing

The fourth phase is where all analyses will take place: send the code, receive data, and begin completing analyses. This phase often takes the longest, as studies will vary on response time, both initially and throughout the life of the project, and issues (e.g., code

errors, convergence issues, assumption violations) will come up. For the datasets that cannot be shared, send detailed instructions describing the exact data requirements, and well-annotated and automated syntax so the analyst can run the cleaning scripts and models with minimal input, issues, errors, or divergence from the other studies. In cases where modelling decisions require knowledge of the data itself (e.g., in multi-state models), the project leader can still automate the scripts to the extent possible and pre-register the decision criteria, but the analytic process will proceed in multiple stages and subsequent scripts tailored based on the results of the first stages. Although as much of the analysis will be automated as possible, each output still requires careful human attention and quality checking for errors or other unforeseen problems.

During this phase, to the extent that all possible data decisions haven't been effectively planned for and scripted, variations in analyst decisions could leak into the project, and the project leader can lose control of the researcher degrees of freedom. This will impact the comparability and credibility of results (Silberzahn et al., 2018). The project leader should know exactly what data are available in a given dataset, and give specific instructions to the data analysts tailored to their specific study, thereby reducing the risk of unintended deviations on the part of the data manager/analyst. We urge the use of a statistical program that has the flexibility to streamline and automate this process as much as possible, by allowing researchers to update models automatically based on available variables, and output specific parts of an analysis without violating data sharing agreements.

This process of planning for contingencies in advance reduces the number of researcher degrees of freedom once results are known. There is added complexity when planning a coordinated analysis: the problems that arise when steps are taken out of order are amplified, as re-coding, re-transforming, re-analyzing anything will not only affect the project leader, but will require new data requests or new analyses from the individual study sites. While correcting small errors in an analysis may take a short time in a single study project, these corrections can take weeks when collaborating across multiple study sites, especially when they cross international time zones. Once the analyses are finalized and compiled from the various data sources, the researcher will summarize the results according to the prescribed plan.

Example Application of Phase 4: Analysis and Model Summarizing—Of the 15 samples included in this coordinated analysis, six chose to conduct analyses internally rather than to share data. We shared three files with the analysts for each of these six samples: a document describing how to prepare a wide-format .csv file with all of the relevant variables, the standardized data preparation script, and the standardized analysis script. The analysts were carefully guided through the data preparation script and were instructed not to make any changes to the analysis script, to alert us if they encountered problems, and to carefully review their output before sending. We conducted analyses for the remaining nine samples using the same standardized procedure and scripts.

Once we received output files for all of the samples, we used the estimates generated from the output files as input for the meta-analysis script. Because the code was prepared during

Phase 2, this step was fully automated. The script produced a folder containing all of the tables and figures for the manuscript and a complete set of results.

Phase 5: Manuscript Preparation

Finally, when the above phases are mostly complete, the investigative team proceeds with writing up the manuscript for publication. As is the case with registered reports, most of the introduction and methods sections can be written in advance, so this phase will focus on the results and discussion. There will likely be more authors than usual, so researchers should anticipate extra time for the revision process. Using a document sharing platform (e.g. Google Docs) can help with version control, but use a method that works best for all authors. We suggest making a priori decisions about authorship, perhaps using the CREDIT system (VandenBos, 2010).

Application of Phase 5: Manuscript Preparation—Fifth, once the meta-analyzed results are final, we will prepare the manuscript (as of this writing, the example coordinated analysis has not yet reached Phase 5). We drafted the introduction during Phase 1, the method during phases Phase 2 and 3, and the results and discussion will be written in Phase 4. Given the large number of authors involved in this coordinated analysis, we use Google Docs to easily share the manuscript with co-authors and to allow for simultaneous editing.

The Importance of Coordinated Analysis.

Lifespan researchers have long been aware of the need for better replicability and credibility of accumulating knowledge. Major strides have been made ideologically, methodologically, and quantitatively, to improve the quality of the published literature. Coordinated analysis was pioneered and applied in the IALSA research network for advancing replicable longitudinal research and integrative analysis of longitudinal studies of aging (as described in Hofer & Piccinin, 2009). Decades of data collection of numerous high quality independent longitudinal data sets of aging have become a rich resource for studying within person changes across a multitude of domains, and this power needs to be harnessed systematically in order to put forth robust and authoritative findings about developmental processes across adulthood (see also: Hofer & Alwin, 2008; Hofer & Sliwinski, 2006; Piccinin & Hofer, 2008). As a result of these developments, integrative analysis has gained in popularity over the last decade in the psychological sciences.

This approach is particularly important for research areas that rely on difficult to obtain data, such as adult development and aging, which face a unique challenge when approaching replicability and open research practices. On the one hand, longitudinal research suffers less from certain issues. For example, most longitudinal studies have large samples and multiple waves of measurement allowing for reliable estimates of change. On the other hand, investigations that use these studies are not immune to questionable research practices such as capitalizing on chance, flexible data analysis, unplanned data-dependent analytic decisions, overfitting models, and selection bias (Nelson et al., 2018; Simmons, Nelson, & Simonsohn, 2011). Coordinated analysis can help reduce the chances of reporting false positives by creating a set of analyses using multiple datasets that are planned and pre-registered in advance, and using, where possible, the same code and models across

all datasets. Results from any one individual sample could be a false positive, false negative, true positive or true negative but when coordinated, consistency across the bulk of the samples will point to both the most probable "true" association as well to possible exceptions due to characteristics of the sample or to Type I or Type II errors. It is also always possible that any consistency is due to an alternative explanation either data or model dependent. Coordinated analysis allows researchers to shift focus from only null hypothesis significance testing (NHST), to deeper examinations of effect sizes, and the consistency of the magnitude and direction of effects across studies. Evaluation of heterogeneity is both a first and a last step in coordinated analysis, with consideration of study-level characteristics to expect and to explain potential variation in effect sizes. In this way, coordinated analysis can help bring longitudinal research into the modern open science and transparency culture. At minimum, it can enhance the rigor and accountability of longitudinal work.

By keeping all datasets separate, coordinated analysis relaxes the strict measurement equivalence that is needed for other forms of IDA such as pooled analysis (Curran & Hussong, 2009), and embraces the idea that there may be a range of true effects, and this is addressed by the random effects meta-analytic model (Borenstein et al., 2010). Since it is likely that the true effect differs from sample to sample, the random effects model is most appropriate for coordinated analysis because RMA explicitly addresses the uncertainty introduced when studies are not identical (Hedges et al., 1998). Alignment of constructs is required, but constructs may be measured differently and have different psychometric properties across studies. Coordinated analysis relies on the assumption that, however coded, the underlying construct is the same across studies. This accommodates the reality of psychological research, and many other kinds of research, in which studies started at different times or in different countries, or may have used different instruments to measure a construct. This of course will introduce noise around the signal, and increase heterogeneity of effects across studies, but all of this can be addressed within the context of coordinated analysis. While harmonization at the construct level in a coordinated analysis may introduce error due to measurement variability into the model, it also begins to take steps toward the issue of generalizability. Coordinated analysis preserves the heterogeneity of studies. Such variation, under the random effects model (Borenstein et al., 2010) is presumed to be due to two main sources: within sample estimation error variance and between study variance. The latter may be caused by variation in all manner of study characteristic differences, such as sampling method, selection effects, cohort differences, start year of study studies (e.g., period effects), country of origin, and of course differences in construct measurement.

If a finding is consistent across the myriad differences in sampling and measurement, then we can say something stronger than "this finding replicates." We can begin to say "this finding generalizes." Obtaining similar results in spite of different sample and measurement characteristics begins to move across the spectrum from replicability to generalizability (for a more detailed discussion of this, see (Condon et al., 2017). The inherent risk in a coordination effort involving variations in design, samples, and measurement is that it can be difficult to identify the source of differences when results are not similar across studies. Is the heterogeneity due to sampling variance or to substantive differences across studies? In the latter case, this means a given effect may be stronger or weaker or even non-existent as one moves across different cultures, historical years of a study, number of waves of

measurement, specific measures of a construct, modes of measurement (e.g., self-report vs, other-report), and age ranges of participants, among many other factors. When coordinated analyses include samples that vary widely on many characteristics, meta-analytic summaries may help evaluate the extent of heterogeneity and permit some accounting, though this is limited by the number of studies available with particular characteristics. When few studies share a particular characteristic in a given coordinated analysis, the study-level power will be low and the meta-analytic summaries should therefore be interpreted with caution and paired with narrative descriptives of the similarities or differences across studies in the magnitude and direction of effects.

While traditional meta-analysis relies on the published literature, attempts are often made to identify file drawer studies so as to minimize publication bias. Coordinated analysis can minimize publication bias because it does not rely on previously published work, which historically has prioritized statistically significant results. However, depending on the breadth of the datasets available for a coordinated analysis, we cannot assume such analyses are free of bias. That said, the fact that 1) multiple datasets are brought to bear on a research question, 2) non-statistically significant effect sizes are included, and 3) an attempt is made to both evaluate the heterogeneity of results and summarize them (often meta-analytically), makes coordinated analysis a form of IDA that is unique in its capacity to promote cumulative scientific knowledge.

Limitations of Coordinated Analysis

There will inevitably be circumstances under which even conceptual harmonization is not possible, for example if the questions require highly specialized data structures, sensitive data, or constructs that are not commonly measured. In some of these cases, coordinated analysis may not be appropriate or even possible. Additionally, even when using the same scale across many datasets, interpretations will be different depending on the context in which it was used, for example across different cultures, cohorts, or historical periods. These issues are not always insurmountable, and indeed can be an asset when handled appropriately, but do require mindful decision making and transparency. It is also important to keep in mind that, with any analysis, including coordinated analysis, differences in modelling decisions will have potential impacts on the output obtained and the conclusions drawn (Marroig et al., 2019).

Lastly, and possibly the most important limitation of coordinated analysis given current incentive structures, is the potential impact on early career researchers (ECRs). While coordinated analysis is, overall, a way to potentially accelerate scientific progress, individual projects and manuscripts are much more time intensive than those focused on a single dataset. The incentive structure of our current academic system requires scholars to publish at a high rate in order to be successful, so adopting coordinated analysis as a common practice may not be practical for ECRs. And yet, we believe that young academics are indeed the best suited for adopting these methods, for a number of reasons. First, these scholars are typically at the top of their analytic skill set, and have been recently trained in state of the art quantitative techniques and programming languages. These skills are essential to scientific advancement and improvement. Second, ECRs aren't bogged down by

the pressure to preserve a theoretical legacy, and may approach the planning of a coordinated analysis with a more open mind and less propensity towards confirmation bias. For these two reasons, we believe that encouraging ECRs to adopt coordinated analysis in their research programs is of utmost importance, and we call on senior colleagues to support these efforts. In general, as a field, we need to become more collaborative, and to normalize multi-author work. This can be accomplished by deliberately pulling in ECRs to collaborate on senior-investigator led projects (e.g., Cadar et al., 2017) or by conducting workshops to complete analyses for coordinated analysis projects (e.g., Duggan et al., 2019; Graham et al., 2020b; Robitaille et al., 2018; Turiano et al., 2020; Weston et al., 2020; Zammit et al., 2021). What's more, our academic institutions would do well to reward the publication of high impact papers in which an ECR is not the lead or senior author. That said, the current institutional incentive structures are unlikely to change in the immediate future, and so we encourage researchers to adopt this approach with a collaborative mindset in order to maximize the number of publications while preserving the integrity and pace of coordinated analysis.

Data Networks and Repositories

As described above, coordinated analysis can be particularly useful to researchers who rely on large long-term longitudinal datasets. Quite a few longitudinal studies make their data publicly available via a variety of platforms (e.g., Inter-University Consortium for Political and Social Research [ICPSR], and National Archive of Computerized Data on Aging [NACDA]), but reliance on these public data sources could introduce new bias into the resulting publications and constrain the generality of findings because many existing longitudinal datasets are not public. For various reasons (e.g., the laws of specific countries, the regulations of certain university systems), making some data public is not possible. Yet some of these data are accessible to qualified researchers even if they cannot be made public.

The Integrative Analysis of Longitudinal Studies of Aging and Dementia (IALSA), is an aging and lifespan-oriented research network that promotes access to both public and non-public longitudinal studies from around the world. Many of the participating studies are not publicly accessible, but the principal investigators have agreed to set up project-specific data sharing agreements or data analysts to help accelerate the publication of replicable longitudinal findings. A public metadata repository and variable searching tool is available through Maelstrom Research (https://www.maelstrom-research.org/network/ialsa) for anyone to search simultaneously for datasets containing the key variables one might need to address their research questions (Fortier, Doiron, Wolfson, & Raina, 2012; Fortier et al., 2016; Hofer & Piccinin, 2009). IALSA demonstrates a collaborative framework and provides a central resource for researchers to gain information on multiple sources of data and streamline the planning phases of a coordinated analysis in a way that was previously much more time consuming.

Data sharing networks help streamline the process of initiating a coordinated analysis, thereby increasing the rigor, credibility, replicability and cumulative nature of longitudinal research findings. Wider use of these networks will also foster a culture of inclusivity as scholars who may not otherwise have collaborated find themselves cooperating and

authoring manuscripts together. This kind of collaboration improves the trustworthiness of network products, mainly because the multi-phase, multi-author process is highly inclusive and reduces the influence of investigator bias, which in single-lab, single-study investigations can become a serious problem. Another benefit of network data use is that it can provide an excellent context through which to train early career researchers. For example, any individual study has limited resources, but the burdens associated with many data sharing requests can be partially alleviated by identifying early career researchers to lead-author coordinated analysis projects: they will not only benefit from collaboration and networking with others in their field, they will also benefit from potentially higher impact papers than would otherwise be the case.

Conclusion

Coordinated analysis is a potentially powerful approach to accumulating knowledge in the study of aging that provides a window into generalizability and external validity. Coordinated analyses are easier to conduct than they have been in the past, due to the establishment of data networks, and public resources such as searchable data archives and measurement catalogues. Additionally, coordinated analysis relies on and promotes central features of efforts to improve open science including, for example, the production of dataset-independent workflows, the publication of both positive and null results, and the pre-registration of agreed-upon cross-study coordinated analysis plans in centralized repositories. Moving forward, greater use of these networks and resources can help accelerate the accumulation of knowledge in lifespan development and other areas that use longitudinal data. This will decrease the amount of uncertainty surrounding bodies of scientific findings in these areas. The possibilities for accelerated, replicable, and generalizable scientific progress facilitated by coordinated analysis are exciting, and are bound only by creativity and our ability to collaborate.

Acknowledgments

References

Bendayan R, Cooper R, Wloch EG, Hofer SM, Piccinin AM, & Muniz-Terrera G (2017). Hierarchy and speed of loss in physical functioning: A comparison across older us and english men and women. Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences, 72(8), 1117–1122.

Bendayan R, Kelly A, Hofer SM, Piccinin AM, & Muniz-Terrera G (2020). Memory decline and depression onset in us and european older adults. Journal of Aging and Health, 32(3–4), 189–198. [PubMed: 30466361]

Borenstein M, Hedges L, Higgins J, & Rothstein H (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. res synth methods 1: 97–111. [PubMed: 26061376]

Brown CL, Gibbons LE, Kennison RF, Robitaille A, Lindwall M, Mitchell MB, … others. (2012). Social activity and cognitive functioning over time: A coordinated analysis of four longitudinal studies. Journal of Aging Research, 2012.

Cadar D, Robitaille A, Clouston S, Hofer SM, Piccinin AM, & Muniz-Terrera G (2017). An international evaluation of cognitive reserve and memory changes in early old age in 10 european countries. Neuroepidemiology, 48(1–2), 9–20. [PubMed: 28219074]

Cadar D, Stephan BC, Jagger C, Johansson B, Hofer SM, Piccinin AM, & Muniz-Terrera G (2016). The role of cognitive reserve on terminal decline: A cross-cohort analysis from two european studies: OCTO-twin, sweden, and newcastle 85+, uk. International Journal of Geriatric Psychiatry, 31(6), 601–610. [PubMed: 26471722]

Chambers CD, Dienes Z, McIntosh RD, Rotshtein P, & Willmes K (2015). Registered Reports: Realigning incentives in scientific publishing. Cortex, 66, A1–A2. 10.1016/j.cortex.2015.03.022 [PubMed: 25892410]

Condon DM, Graham EK, & Mroczek DK (2017). On Replication Research. PsyArXiv.

Curran PJ, & Hussong AM (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. Psychological Methods, 14(2), 81. 10.1037/a0015914 [PubMed: 19485623]

Duggan EC, Piccinin AM, Clouston S, Koval AV, Robitaille A, Zammit AR, … others. (2019). A multi-study coordinated meta-analysis of pulmonary function and cognition in aging. The Journals of Gerontology: Series A, 74(11), 1793–1804. 10.1093/gerona/glz057

Fortier I, Doiron D, Wolfson C, & Raina P (2012). Harmonizing Data for Collaborative Research on Aging: Why Should We Foster Such an Agenda? Canadian Journal on Aging / La Revue Canadienne Du Vieillissement, 31(01), 95–99.

Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, … Burton P (2016). Maelstrom Research guidelines for rigorous retrospective data harmonization. International Journal of Epidemiology, 66, dyw075. 10.1093/ije/dyw075

Ghersi D, Berlin J, & Askie L (2011). Cochrane prospective meta-analysis methods group. COCHRANE METHODS, 35.

Graham EK, Rutsohn JP, Turiano NA, Bendayan R, Batterham PJ, Gerstorf D, … Mroczek DK (2017). Personality predicts mortality risk: An integrative data analysis of 15 international longitudinal studies. Journal of Research in Personality, 70, 174–186. 10.1016/j.jrp.2017.07.005 [PubMed: 29230075]

Graham EK, Weston SJ, Gerstorf D, Yoneda TB, Booth T, Beam CR, … others. (2020a). Trajectories of big five personality traits: A coordinated analysis of 16 longitudinal samples. European Journal of Personality. 10.1002/per.2259

Graham EK, Weston SJ, Turiano N, Aschwanden D, Booth T, Harrison F, … others. (2020b). Is healthy neuroticism associated with health behaviors? A coordinated integrative data analysis. Collabra: Psychology, 6(1), 32. 10.1525/collabra.266 [PubMed: 33354649]

Griffith LE, Van Den Heuvel E, Fortier I, Sohel N, Hofer SM, Payette H, … others. (2015). Statistical approaches to harmonize data on cognitive measures in systematic reviews are rarely reported. Journal of Clinical Epidemiology, 68(2), 154–162. [PubMed: 25497980]

Griffith L, Heuvel E. van den, Fortier I, Hofer SM, Raina P, Sohel N, … Belleville S (2013). Harmonization of cognitive measures in individual participant data and aggregate data meta-analysis.

Hedges L, & Vevea JL (1998). Fixed-and random-effects models in meta-analysis. Psychological Methods, 3(4), 486.

Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG. Meta-analysis of continuous outcome data from individual patients. Stat Med. 2001 Aug 15;20(15):2219–41. doi: 10.1002/sim.918. [PubMed: 11468761]

Hofer SM, & Alwin DF (2008). Handbook of cognitive aging: Interdisciplinary perspectives. Sage.

Hofer SM, & Piccinin AM (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. Psychological Methods, 14(2), 150–164. 10.1037/a0015566 [PubMed: 19485626]

Hofer SM, & Piccinin AM (2010). Toward an integrative science of life-span development and aging. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 65(3), 269–278. 10.1093/geronb/gbq017

Hofer SM, & Sliwinski MJ (2006). Design and analysis of longitudinal studies of aging. In Birren JE & Schaie KW (Eds.), Handbook of the psychology of aging (6th Edition; pp. 15 37). San Diego: Academic Press.

Hoogendijk EO, Rijnhart JJ, Skoog J, Robitaille A, Hout A. van den, Ferrucci L, … others. (2020). Gait speed as predictor of transition into cognitive impairment: Findings from three longitudinal studies on aging. Experimental Gerontology, 129, 110783. [PubMed: 31751664]

Kelly A, Calamia M, Koval A, Terrera GM, Piccinin AM, Clouston S, … Hofer SM (2016). Independent and interactive impacts of hypertension and diabetes mellitus on verbal memory: A coordinated analysis of longitudinal data from england, sweden, and the united states. Psychology and Aging, 31(3), 262. [PubMed: 26913926]

Lindwall M, Cimino CR, Gibbons LE, Mitchell MB, Benitez A, Brown CL, … others. (2012). Dynamic associations of change in physical activity and change in cognitive function: Coordinated analyses of four longitudinal studies. Journal of Aging Research, 2012.

Marroig A, uki I, Robitaille A, Piccinin AM, & Terrera GM (2019). Importance of modelling decisions on estimating trajectories of depressive symptoms and co-morbid conditions in older adults: Longitudinal studies from ten European countries. PloS one, 14(4), e0214438. [PubMed: 30943221]

Mitchell MB, Cimino CR, Benitez A, Brown CL, Gibbons LE, Kennison RF, … others. (2012). Cognitively stimulating activities: Effects on cognition across four studies with up to 21 years of longitudinal data. Journal of Aging Research, 2012.

Mroczek DK, & Spiro A (2007). Personality change influences mortality in older men. Psychological Science, 18(5), 371–376. 10.1111/j.1467-9280.2007.01907.x [PubMed: 17576273]

Nelson LD, Simmons JP, & Simonsohn U (2018). Psychology's Renaissance. Annual Review of Psychology, 69(1), 511–534. 10.1146/annurev-psych-122216-011836

Patsopoulos NA, Evangelou E, & Ioannidis JP (2008). Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. International Journal of Epidemiology, 37(5), 1148–1157. 10.1093/ije/dyn065 [PubMed: 18424475]

Piccinin AM, & Hofer SM (2008). Integrative analysis of longitudinal studies on aging: Collaborative research networks, meta-analysis, and optimizing future studies. 10.4135/9781412976589.n27

Piccinin AM, Muniz G, Sparks C, & Bontempo DE (2011). An evaluation of analytical approaches for understanding change in cognition in the context of aging and health. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 66(suppl_1), i36–i49.

Piccinin AM, Muniz-Terrera G, Clouston S, Reynolds CA, Thorvaldsson V, Deary IJ, … others. (2013a). Coordinated analysis of age, sex, and education effects on change in mmse scores. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68(3), 374–390. 10.1093/geronb/gbs077

Robitaille A, Hout A. van den, Machado RJ, Bennett DA, uki I, Deary IJ, … others. (2018). Transitions across cognitive states and death among older adults in relation to education: A multistate survival model using data from six longitudinal studies. Alzheimer's & Dementia, 14(4), 462–472. 10.1016/j.jalz.2017.10.003

Rodgers JL, & Shrout PE (2018). Psychology's replication crisis as scientific opportunity: A précis for policymakers. Policy Insights from the Behavioral and Brain Sciences, 5(1), 134–141. 10.1177/2372732217749254

Scott SB, Sliwinski MJ, Zawadzki M, Stawski RS, Kim J, Marcusson-Clavertz D, Lanza ST, Conroy DE, Buxton O, Almeida DM, Smyth JM. A Coordinated Analysis of Variance in Affect in Daily Life. Assessment. 2020 Dec;27(8):1683–1698. doi: 10.1177/1073191118799460. [PubMed: 30198310]

Sharp ES, Beam CR, Reynolds CA, & Gatz M (2019). Openness declines in advance of death in late adulthood. Psychology and Aging, 34(1), 124. 10.1037/pag0000328 [PubMed: 30667240]

Silberzahn R, Uhlmann EL, Martin DP, Anselmi P, Aust F, Awtrey E, … Nosek BA (2018). Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. Advances in Methods and Practices in Psychological Science, 67, 337–356. 10.1177/2515245917747646

Simmons JP, Nelson LD, & Simonsohn U (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science, 22(11), 1359–1366. 10.1177/0956797611417632 [PubMed: 22006061]

Stawski RS, Scott SB, Zawadzki MJ, Sliwinski MJ, Marcusson-Clavertz D, Kim J, … Smyth JM (2019). Age differences in everyday stressor-related negative affect: A coordinated analysis. Psychology and Aging, 34(1), 91. [PubMed: 30550311]

Turiano N, Graham EK, Weston SJ, Booth T, Harrison F, Lewis N, … others. (2020). Is healthy neuroticism associated with longevity? A coordinated integrative data analysis. Collabra: Psychology, 6(1), 33. 10.1525/collabra.268 [PubMed: 33354648]

Van den Akker O, Weston SJ, Campbell L, Chopik WJ, Damian RI, Davis-Kean P, … others. (2019). Preregistration of secondary data analysis: A template and tutorial.

VandenBos GR (Ed.). (2010). Publication manual of the American Psychological Association (6th ed.). Washington, D.C.: American Psychological Association.

Vazire S (2018). Implications of the credibility revolution for productivity, creativity, and progress. Perspectives on Psychological Science, 13(4), 411–417. 10.1177/1745691617751884 [PubMed: 29961410]

Weston SJ, Graham EK, & Piccinin A (2019). Coordinated data analysis: A new method for the study of personality and health. In Hill P & Allemand M (Eds.), Personality and healthy aging in adulthood. Springer Nature.

Weston SJ, Ritchie SJ, Rohrer JM, & Przybylski AK (2019). Recommendations for increasing the transparency of analysis of preexisting data sets. Advances in Methods and Practices in Psychological Science, 2, 214–227. [PubMed: 32190814]

Weston SJ, Graham EK, Turiano N, Aschwanden D, Booth T, Harrison F, … others. (2020). Is healthy neuroticism associated with chronic conditions? A coordinated integrative data analysis. Collabra: Psychology, 6(1), 42. 10.1525/collabra.267 [PubMed: 33073161]

Willroth E, Beck E, Beam C, Booth T, Drewelies J, Gerstorf D, … Graham E (in prep). Does personality change predict mortality risk?: A coordinated analysis of 15 longitudinal studies.

Willroth E, Graham EK, & Mroczek DK (in press). Challenges and opportunities in pre-registration of coordinated data analysis: A tutorial and template. Psychology and Aging.

Wood AM, Kaptoge S, Butterworth AS, Willeit P, Warnakula S, Bolton T, … others. (2018). Risk thresholds for alcohol consumption: Combined analysis of individual-participant data for 599 912 current drinkers in 83 prospective studies. The Lancet, 391(10129), 1513–1523. 10.1016/S0140-6736(18)30134-X

Yoneda TB, Rush J, Graham EK, Berg AI, Comijs H, Katz M, … Piccinin AM (2020a). Increases in Neuroticism May Be an Early Indicator of Dementia: A Coordinated Analysis. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 75, 251–262. 10.1093/geronb/gby034 [PubMed: 29608748]

Yoneda T, Lewis NA, Knight JE, Rush J, Vendittelli R, Kleineidam L, Hyun J, Piccinin AM, Hofer SM, Hoogendijk EO, Derby CA, Scherer M, Riedel-Heller S, Wagner M, van den Hout A, Wang W, Bennett DA & Terrera GM (2020b [advance access]). The Importance of Engaging in Physical Activity in Older Adulthood for Transitions between Cognitive Statues Categories and Death: A Coordinated Analysis of Fourteen Longitudinal Studies. The Journals of Gerontology: Medical Sciences.

Zammit AR, Piccinin AM, Duggan EC, Koval A, Clouston S, Robitaille A, … others. (2021). A coordinated multi-study analysis of the longitudinal association between handgrip strength and cognitive function in older adults. The Journals of Gerontology: Series B, 76(2):229–241. doi: 10.1093/geronb/gbz072.
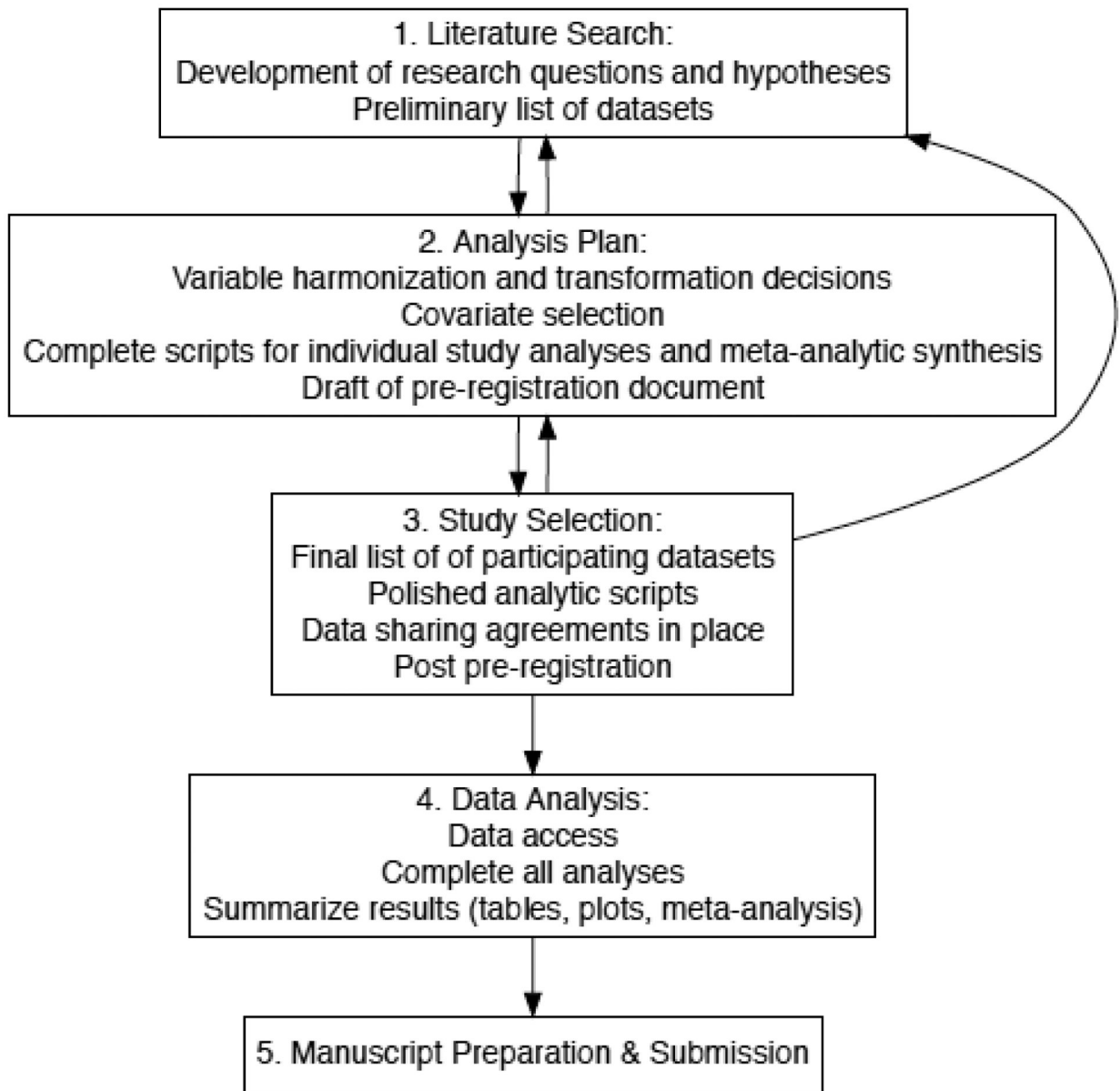
Figure 1.
Flowchart of the typical phases in a coordinated analysis.