

Meta-analysis for Sociology – A Measure-driven Approach

Bulletin de Méthodologie Sociologique
117 75–92

© The Author(s) 2013

Reprints and permission:

sagepub.com/journalsPermissions.nav

DOI: 10.1177/0759106312465554

bms.sagepub.com



David J. Roelfs

University of Louisville, KY, USA

Eran Shor

McGill University, Quebec, Canada

Louise Falzon

Columbia University, NY, USA

Karina W. Davidson

Columbia University, NY, USA

Joseph E. Schwartz

Columbia University & Stony Brook University, NY, USA

Résumé

Méta-analyse pour la sociologie - Une approche orientée « mesure » : Les méthodes de méta-analyse ont pris de l'importance dans la recherche sociologique. Dans cet article, nous présentons une approche de méta-analyse spécialement utile aux sociologues. Les approches conventionnelles de la méta-analyse donnent souvent priorité à l'analyse de la bibliographie du champ fondée sur une approche orientée « concept ». Or, dans les disciplines ayant une grande diversité de théories, comme la sociologie, cette approche contraint le chercheur à exploiter complètement et en totalité la production du domaine étudié. Nous expliquons qu'une approche orientée « mesure », au cours de laquelle des recherches itératives et de nouvelles techniques de recherche automatisées sont utilisées permet d'augmenter le nombre de publications trouvées (et

Corresponding Author:

Eran Shor, Department of Sociology, McGill University, 855 Sherbrooke Street West, Montreal, Quebec H3A 2T7, Canada.

Email: ershor@gmail.com

ainsi l'éventail des analyses possibles), d'élargir la période considérée et de d'outrepasser les frontières des disciplines. Nous décrivons cette approche orientée « mesure » à partir de deux projets de méta-analyse qui examinent les effets de différentes variables sociales sur toutes les causes de mortalités.

Abstract

Meta-analytic methods are becoming increasingly important in sociological research. In this article we present an approach for meta-analysis which is especially helpful for sociologists. Conventional approaches to meta-analysis often prioritize “concept-driven” literature searches. However, in disciplines with high theoretical diversity, such as sociology, this search approach might constrain the researcher’s ability to fully exploit the entire body of relevant work. We explicate a “measure-driven” approach, in which iterative searches and new computerized search techniques are used to increase the range of publications found (and thus the range of possible analyses) and to traverse time and disciplinary boundaries. We demonstrate this measure-driven search approach with two meta-analytic projects, examining the effects of various social variables on all-cause mortality.

Mots clés

Méta-analyse, Sociologie, Orientée « mesure », Recherches alternatives, Causes de mortalité

Keywords

Meta-analysis, Sociology, Measure-driven, Alternative Search, Mortality

The use of meta-analysis¹ in sociology has been rapidly growing in recent years. Sociologists have often adopted other disciplines’ literature search conventions (which we label here “concept-driven” approaches) that rely on a tight connection between the concepts being studied and the keywords used to classify research reports. The “ideal type” concept-driven search model is comprised of two main stages. First, sources are identified using a keyword search in bibliographic data bases, complemented with searches in literature reviews, consultations with experts, and searches of highly-relevant journals. The researcher then evaluates this initial set of publications for study eligibility. Second, additional sources are identified through searches of the bibliographies of coded publications. Adherence to a concept-driven approach lowers the dangers of introducing heterogeneity (see Huedo-Medina et al., 2006), and also rapidly produces closure, suggesting that the literature has been exhausted. While researchers often adopt a broader approach, an examination of 20 randomly selected meta-analyses shows that current methodological practice and reporting, perhaps inadvertently, encourage the use of a concept-driven approach in all research settings (see Table S1 in the Online Supplement and the associated discussion for an expanded explanation of this point).

In some disciplines such as clinical psychology and medicine, the reliance on narrowly-defined search terminology is less problematic as there is more widespread

agreement on which theoretical concept (keyword) a particular measure represents. In sociology and other social sciences, however, scholars often disagree on the best way to operationalize concepts such as gender inequality, socio-economic status, civil society, terrorism, or organizational success. As a result, one can find various empirical measures for each of these concepts in the literature. To further complicate matters, many of these various empirical measures are sufficiently general so as to serve as an adequate measure for another concept from an entirely separate sub-field of study. For example, the concept of gender inequality can be measured in myriad ways, including through gendered comparisons of wages and promotions, domestic violence occurrences, or time allocated to housework. At the same time, a different study seeking to measure how much a person is “overworked” might also examine time allocations and therefore include “time spent doing housework” (by gender) as a covariate without necessarily ever addressing gender issues. The multiplicity of sub-fields and disciplines using a given measure creates inconsistencies in the way social science research is indexed (see Becker and Sanders 2006), which in turn renders the concept-driven approach less suitable.

While many sociologists, including those conducting meta-analyses, are aware of this problem, it has not been thoroughly addressed by the literature on meta-analytic methods. We therefore seek to explicate an alternative “ideal type”: a “measure-driven” approach. This approach draws from the insights of qualitative sociological methods regarding the benefits of approaching one’s research with an open mind and using the data to refine the subsequent analysis and search procedures; see Glaser and Strauss (1967), Strauss and Corbin (1990, 1994), and Burawoy (1998). As with the concept-driven approach, the research begins with the design of a keyword search algorithm by a researcher who is well-versed in the literature of his/her sub-field, able to clearly specify the concept being studied, and able to generate a thorough list of the alternative measures of this concept used within the sub-field. Next, however, the researcher uses an iterative search process to exhaust the literature of multiple sub-fields rather than that of the researcher’s sub-field alone. Each search iteration through bibliographies, citing articles, and similar articles constitutes what we call a search “wave.” Search waves are repeated until no additional sources emerge. With each wave, the search results are carefully scrutinized in an effort to capture alternative concepts of which the researchers may have been initially unaware. Our search method then utilizes this expanded theoretical framework to guide subsequent literature searches. In other words, the measure-driven approach encourages eligibility criteria which are organically developed rather than a priori conceived. It is therefore more consistent with the Cochrane Collaboration (Higgins and Green, 2008) and Campbell Collaboration (Farrington and Petrosino, 2001) calls for balance between data base and manual search when conducting meta-analyses and with the growing recognition that meta-analysis is “an iterative process” (Moher et al., 2009).

The concept-driven and measure-driven approaches represent “ideal types” placed at opposing ends of a continuum of meta-analytic practice.² To assess the usefulness of a more measure-driven approach, we present evidence derived from two large projects (each consisting of multiple meta-analyses) that examine the effects of diverse factors on all-cause mortality: The first project examines the effects of various stressful life

events – war stress, unemployment, marital dissolution, and the death of a close person – while the second examines the effects of various social relationship measures – social network size, social contact, social participation, and social support. We compare these two projects to previous similar research endeavors that employed a more concept-driven search approach.

Our comparison highlights several major advantages of the measure-driven approach in a sociological context. First, this approach is much more thorough (it yielded substantially larger numbers of coded publications than the concept-driven approach). Our search results demonstrate that, even if terms are very well-specified, relying on the keywords used within a single sub-field identifies only a fraction of the relevant literature. Second, although the measure-driven approach is more time-consuming, the results it yields may justify the additional effort. Using it, researchers are more likely to locate quality publications from the formative days of a given field and gain access to bodies of literature in multiple disciplines, both of which are useful in reducing selection bias (a problem emphasized by Egger and Davey-Smith (1998), Berman and Parker (2002), Lipsey and Wilson (2001), and Bailar (1997)). Finally, gathering meta-analytic data using a more measure-driven approach has the potential to facilitate not only theory testing, but also theory building.

Meta-analysis Projects

In this section we present two recently completed, extensive meta-analysis projects conducted according to the principles of the measure-driven approach. Each of these projects yielded multiple, distinct meta-analyses rather than one overarching meta-analysis. The first project examines events such as unemployment, war stress and widowhood (all narrowly defined subsets of the overarching theoretical construct “stressful life events”) on all-cause mortality. The second project is based on the body of research that looks at the effects of various social relationships on all-cause mortality. We use these studies to demonstrate the advantages of a measure-driven approach. First, we show the difference between our actual search yield and the yield we would have gotten had we used a more concept-driven approach. Second, we compare our results to those of two previous meta-analyses in the field and show that the measure-driven approach helped produce much more detailed analyses.

One of the primary advantages of the measure-driven approach is its exhaustiveness. In both of our projects, it produced many additional publications of high quality and substantive importance. Figures 1 and 2 show the number of publications coded in each search wave for the stress and social relationships studies respectively. In the stress study (Figure 1), we conducted nine search waves, with the working definition and set of search terms being revisited and revised following each wave. The number of publications we ended up coding rapidly increased, reached a peak in Wave 3, and then exponentially diminished in the subsequent waves. A similar pattern emerged in the social relationships study (Figure 2). Here we conducted five search waves, also reaching a peak in Wave 3, and then a more rapid exponential drop. The striking resemblance of these two figures suggests that this pattern of rise and fall is not a coincidence.

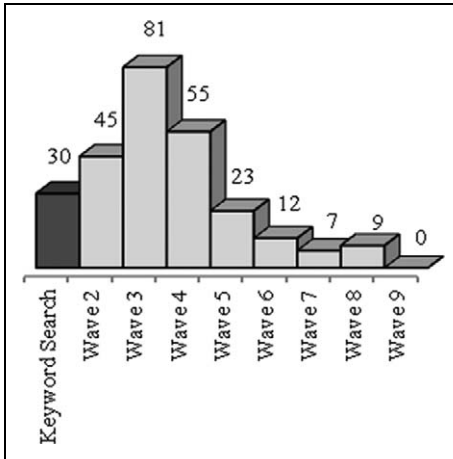


Figure 1. Number of publications coded by search wave for the stressful life events project.

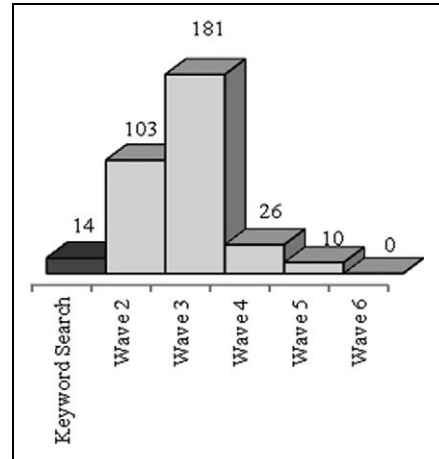


Figure 2. Number of publication coded by search wave for the social relationships project.

Furthermore, both figures show that a substantial number of publications were found only at the later stages of our search.

Project 1 - Stressful Life Events and All-cause Mortality

Search and Coding Procedures

In June 2005, we conducted a keyword search of electronic bibliographic data bases to retrieve all publications combining the concepts of stressful life events (including widowhood, marital dissolution, unemployment, war stress, death of a loved one, experience of a natural disaster, etc.) and all-cause mortality. We used 100 search clauses for Medline, 97 for EMBASE, 81 for CINAHL, and 20 for Web of Science (see the Online Supplement for the full Medline search algorithm). We identified 1,570 unique publications. Using these results as a base, two of the authors iteratively searched the bibliographies of eligible publications, all sources citing an eligible publication, and (at a minimum) 100 publications identified as “similar to” an eligible publication. We exhausted the literature after 8 iterations. The original keyword search was re-run in July 2008 and yielded 910 additional references. All sources produced by this latter search were already identified and considered by the coders in previous searches, suggesting that our search procedure had reached saturation.

Figure 3 summarizes the number of publications considered at each step of the search process. 262 publications were coded and analyzed in this research. We first narrowed down the list of 1,570 publications identified by the original keyword search to 48 publications, 30 of which were eventually coded. In the subsequent eight search waves, we identified 681 publications, 232 of which were also coded. We reviewed an unusually large number of publications (approximately 75,000 titles) identified using publication bibliographies and three online data bases (Web of Knowledge, PubMed, and Google Scholar).

The process of reviewing these 75,000 titles was conducted in three stages. First, due to the large amount of time needed to review each publication's abstract, we began the screening process using only the publication's title. This process of reviewing titles, although still time-consuming, was feasible (we estimate that each of us reviewed approximately 500 titles per hour). Because of the limited amount of information that is contained in the title alone, we adopted an inclusive approach. A publication was excluded at this stage from further consideration only if it was clear from the title that the paper had no potential relevance for our study (the full inclusion/exclusion criteria used are described later). For example, we excluded a publication if the title indicated it was a study of a non-psychosocial variable such as "exposure to radiation and mortality" or "depression and mortality". However, we did not exclude a publication if the title indicated, even very vaguely, that it could have been a study of one or more psychosocial stressors such as "social factors and mortality" or "demographic risk factors and mortality".

In the second stage of the screening process, we carefully reviewed each remaining publication's abstract and tables. Once again, we adopted an inclusive approach. Only publications that clearly did not contain any stressors of interest were excluded from further consideration. For example, if a publication's title was "demographic risk factors and mortality", we would exclude it at this stage from further review if the measurements reported in the abstract and tables were exclusively non-psychosocial – only gender, ethnicity, income, and age. We reviewed the abstracts and tables of 681 publications. In the third and final stage of the screening process, we read each of the 413 remaining publications in their entirety and made the final inclusion/exclusion decisions according to the criteria described below. At the end of this final stage, we were left with 262 publications, all of which were coded.

Inclusion Criteria

To be included in the final analysis, a research report had to provide data on the effects of having experienced one or more stressful life events satisfying the following definition:

A discrete external event (having an identifiable starting point) commonly thought to be mentally stressful, excluding events solely involving a physical illness or injury of the study participant.³

This definition excludes those independent variables corresponding with situations that are merely potentially stressful but where it is unclear if a stressful event actually occurred. For example, several publications examined the effects of firefighting as an occupation on mortality rates because this occupation is potentially stressful. We did not include such studies in this meta-analysis because we had no way of determining whether or not a stressful event actually occurred. For similar reasons, we did not include the several publications that examined the effects of military service, unless the publication also sought to identify those members of the military exposed to stressful events such as combat or capture. The above definition also excludes conditions that are non-discrete – having no concrete starting point – such as never getting married, having a chronic illness, or having a chronic alcohol or drug-abuse problem. Also

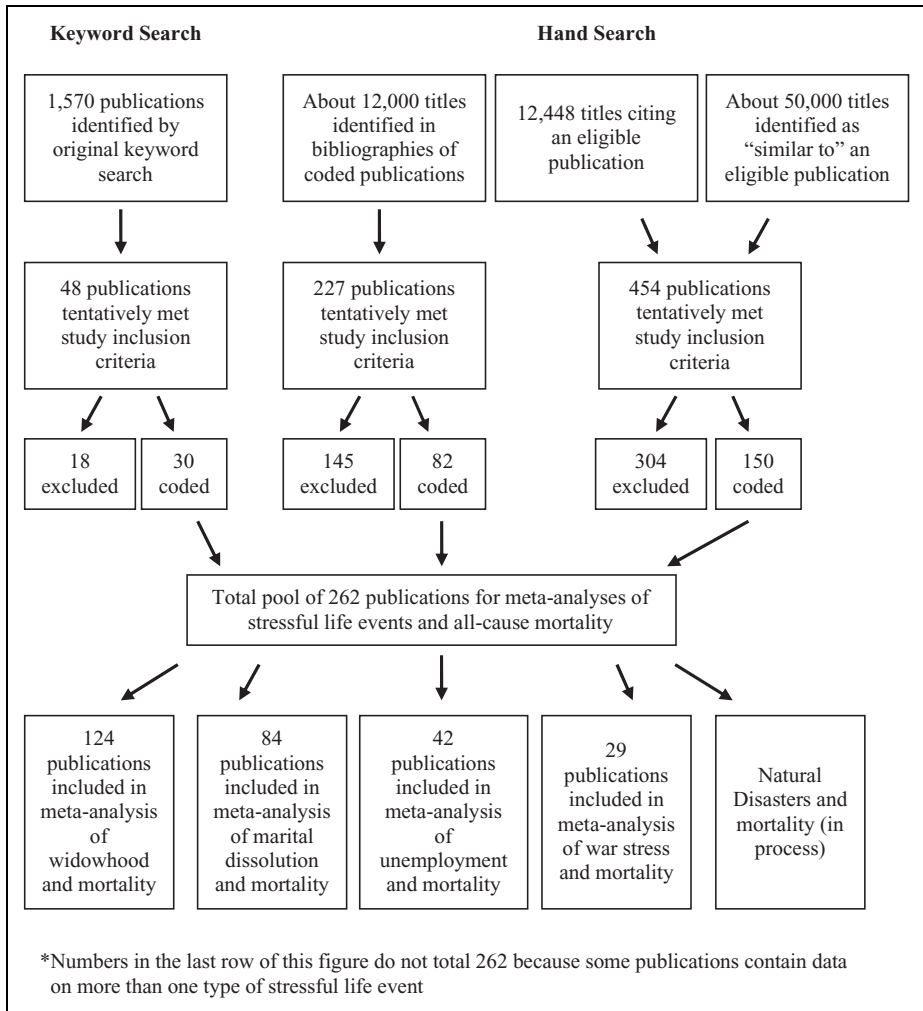


Figure 3. Search strategy and yield for the stressful life events project.

excluded were independent variables measuring the absence of factors thought to mediate stress, such as social support, social integration, or going on vacation. Our definition of a stressful life event also excludes strictly psychological variables such as depression, anxiety, and self-reported stress levels.

The final typology includes the following stressful life events: the death of a loved one, marital dissolution, unemployment, war zone exposure (either as a civilian or as a member of the military), having survived a concentration camp, having been a prisoner of war, having been a refugee, going through a natural disaster, experiencing workplace downsizing, sudden loss of property, involuntary relocation, having a close friend or relative who experienced a highly-stressful event, and having one or more children while single. We also

included discrete financial strains, being careful to exclude independent variables that measured only long-term socioeconomic conditions. Finally, we included any scales that measured a combination of the above stressors.

Next we applied further exclusion criteria to the publications that matched our definition of a stressful life event. We only included studies: (1) reporting all-cause mortality;⁴ (2) making a clear comparison between a group of people that experienced one or more stressful events and another group that either did not experience the same stress at all or experienced it to a lesser degree; and (3) reporting quantitative point estimates and standard errors (or providing some way to calculate them).

Project 2 - Social Relationships and All-cause Mortality

Search and Coding Procedures

The keyword search algorithm used for the stressful life events project also included keywords for social isolation/participation. While these concepts were subsequently excluded from the stress analysis, we coded 14 of the 1,570 original publications for use as a starting point for a second meta-analysis project, focusing on social relationships and their effects on all-cause mortality.

Figure 4 summarizes the number of publications considered at each step of the search process. Using the three-stage screening process described for the previous project, we examined approximately 100,000 publications' titles, 749 abstracts and tables, and read 482 publications in full. Following the final inclusion/exclusion decision, we coded and analyzed a total of 334 publications, 14 of them coming from the original search and 320 from the subsequent four search waves. All other search and coding procedures were done in exactly the same manner described above for the other project.

Inclusion Criteria

To be included in the data base, a research report had to provide data on the effects of one or more of the following social relationships measures on mortality:

1. *Social Contact* – The number of people a respondent has contact with and the frequency with which that contact occurs. This also includes what the literature refers to as social isolation (the absence of social contact).
2. *Social Network Size* – The number of people (such as relatives, friends, and neighbors) with whom the respondent has social ties, regardless of whether or not the respondent actually has contact with them.
3. *Social Participation* – The number of voluntary groups (charity organizations, book clubs, or religious congregations) of which a respondent is a member, and the frequency with which the respondent participates in the activities of these groups.
4. *Emotional Social Support Received* – A subjective evaluation, given by a respondent, of the level of support he or she receives from others.

Several closely-related concepts were excluded from the project. First, we only included measures of emotional social support received, but not support given to others (which is

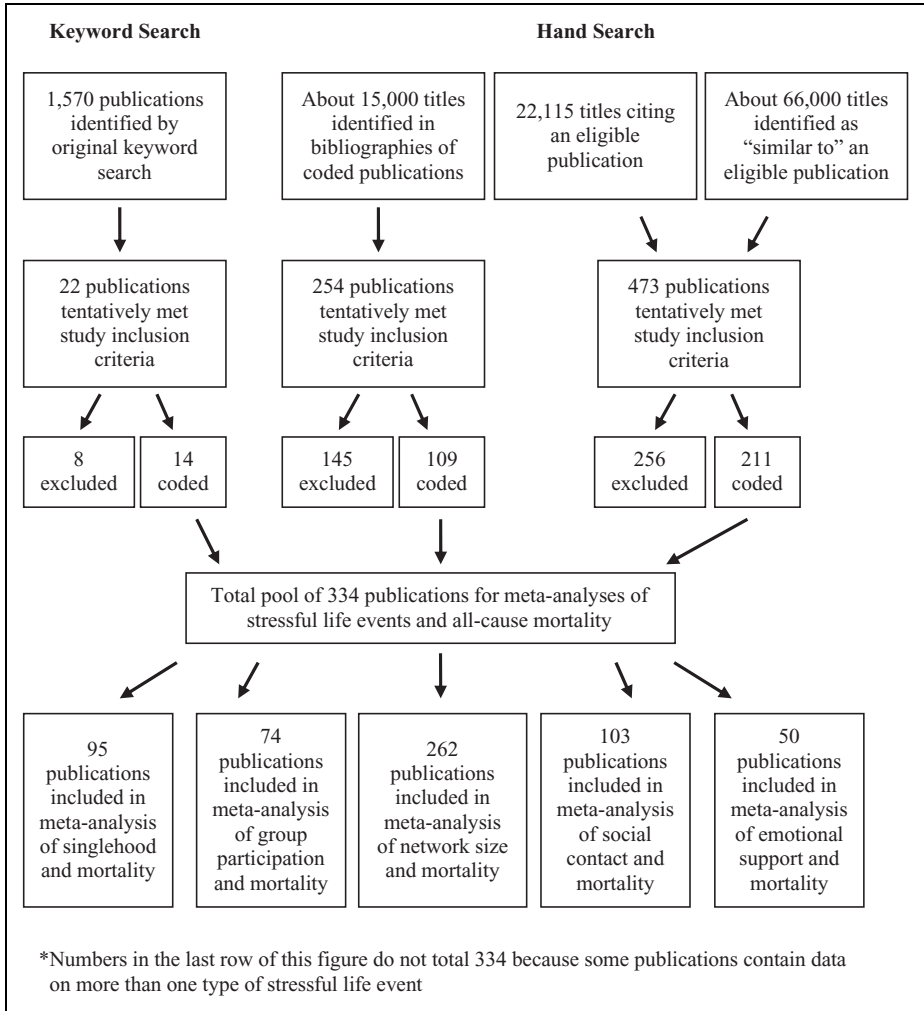


Figure 4. Search strategy and yield for the social relationships project.

often confounded with the respondent’s health and with the ill-health of a close one). Likewise, we did not include measures of instrumental support received, also highly confounded with a respondent’s health status. To keep the focus on voluntary social ties, we excluded contacts with paid help and institutional personnel (home healthcare professionals, government employees, etc.). We also excluded measures of contact or social support highly confounded with stressful life events (the death of a child or confidant, or social isolation as the result of confinement). Finally, we included only measures of clear social contact or participation, thus excluding measures such as attendance at cultural events (movies or concerts) and solitary activities (reading books or gardening).

As with the stressful life events project, we applied additional exclusion criteria. We only included studies: (1) reporting all-cause mortality; (2) making a clear

comparison between a group of people who experienced a high level of social relationships and a group of people who experienced low levels of social relationships; and (3) reporting quantitative point estimates and standard errors (or providing some way to calculate them).

An Empirical Comparison of the Concept-driven and Measure-driven Approaches

The measure-driven approach is quite time-consuming (each wave took an average of three months for two coders, working 10 weekly hours, to complete) and requires a great deal of effort and conceptual flexibility. In a world of limited resources (in terms of both time and money), it is not always realistic to expect researchers to fully follow the procedure we have suggested. Furthermore, some researchers (see Lipsey and Wilson 2001: 24) recommend exercising “restraint” when conducting literature searches in meta-analyses, citing the great labor cost of pursuing “low probability prospects”. However, when conducting the meta-analysis search, sociological researchers should also be aware of the potential advantages of a measure-driven-oriented approach and make a conscious choice regarding the degree to which it should be pursued. In this section, we demonstrate the significant benefits of using such an approach. The section is divided into three parts. First, we examine differences in the search yield between our measure-driven approach and the concept-driven approach. Second, we show that the additional sources found by the measure-driven approach are of high quality. Finally, we describe the analytic leverage provided by obtaining a much larger sample of high-quality publications.

Differences in Search Yield

In Figure 5, we look at each of our case studies, comparing our own initial concept-driven search and our subsequent measure-driven search. The measure-driven search approach yielded approximately six times as many coded publications as the concept-driven search approach (262 versus 46 for the stress meta-analysis, and 334 versus 54 for the social relationships meta-analysis).⁵ This striking departure supports our contention that the measure-driven approach is substantially more thorough.

To further show the differences between the concept-driven and the measure-driven search approaches, Figure 6 compares the yield of two previous meta-analyses to that of our own study. First, McCullough et al. (2000) studied the relationship between religious involvement and all-cause mortality. One of the main measures included in their study was religious attendance, also examined by our own social relationships project. To make sure that the two are comparable, we only report here search results published up to 1999 (employing the same eligibility criteria reported by the previous study). While McCullough et al. found only 13 publications with measures of religious attendance, we found 27 such publications. Second, Manzoli et al. (2007) examined the effects of marital status on all-cause mortality in publications between 1995 and 2005 and among respondents aged 60 years or older. To ensure comparability, we compare these to our search results for the same ages and publication years and using precisely the same eligibility criteria. We found 117 publications,⁶ 12 compared with only 31 found by Manzoli et al. It is clear then that when following

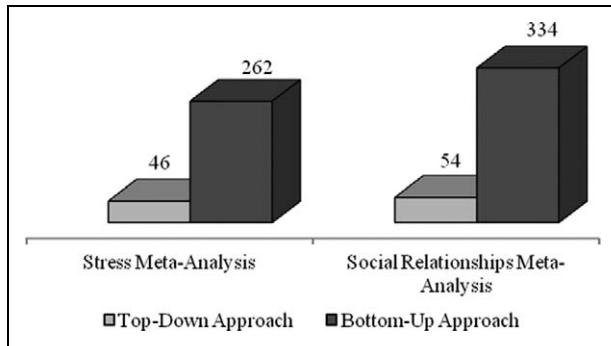


Figure 5. Number of articles coded using the measure-driven search approach compared to the number of articles coded using a concept-driven search approach.

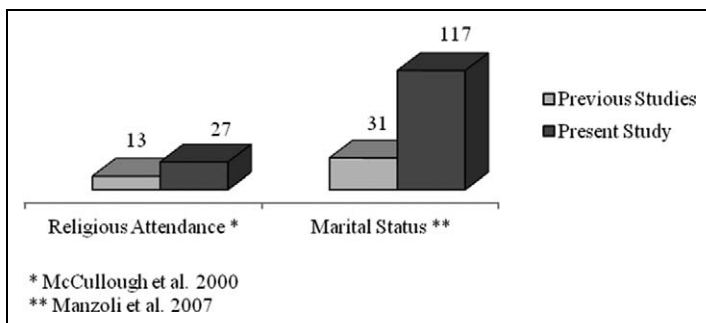


Figure 6. Number of articles coded by previous studies vs. number of articles coded in matching subsets of the present study.

standard meta-analytic practices, even well-conducted searches of narrowly-defined concepts omit a substantial portion of the literature.

Others have also found that keyword searches miss much of the relevant literature. For example, one recent report on meta-analysis search procedures (Greenhalgh and Peacock 2005) found that only 30 percent of coded publications were identified through keyword searches. Our case studies suggest that this problem, as a general rule, is much graver than previously thought. Of the 454 publications coded in the stress project and the social relationships project combined, the Medline, PsycInfo, CINAHL and Embase data bases had records for only 120.⁷ This means that even if one somehow knew all possible keywords from the outset, she/he would have found only about one-fourth of the relevant literature. To illustrate this point, only 2 of our 30 publications on unemployment and 3 of our 27 publications on war stress were among the 120 found in keyword searches of these data bases.

Our results show that the measure-driven approach allowed us to expand the body of coded publications by effectively breaching disciplinary and conceptual boundaries. This is demonstrated by an interesting pattern which emerged in our social relationships project. Our searches showed that the variable “number of children” is a common

operationalization for social network size. The variable “number of children” and its effect on mortality, however, is used not only in studies of social network size, but also as a measure of socioeconomic status (mostly in the sociological literature) and to analyze evolutionary patterns (mostly in the biology and general science literatures). Despite the fact that they are all using the same variable in much the same way, these literatures do not cite each other and often use different terminology. Therefore, in the early search waves (the first two), almost all of the publications we found came from epidemiology and public health journals. In Wave 3, we began to uncover a substantial number of articles in sociology journals. It was not until the last waves that articles from biological and general science journals emerged.

Our use of the measure-driven approach also allowed us to locate more of the literature from the formative days of the field. Scientists often tend to cite publications from recent literature and neglect earlier findings. While this tendency may sometimes be justified with the fact that older publications become methodologically or substantively obsolete, this is problematic as an overarching strategy in meta-analyses, where older publications of high quality must receive full consideration. Furthermore, older publications are less likely to have an abstract in electronic data bases, making it difficult to locate them through a keyword search. In our own study, we found that it was only in the later stages of the search that we were able to locate a substantial number of well-conducted studies from before 1970. This suggests that using the concept-driven approach potentially denies the researcher the opportunity to examine time trends and might introduce selection bias.

Minimizing selection bias is especially important in meta-analysis, because this method claims to summarize the state of the knowledge in a given field. Search approaches in meta-analyses are inherently non-random, as the only way to expand the search is through a snowball procedure. There is no way to guarantee that the collection of publications found by such procedures is representative of the population of interest, but we argue that the measure-driven approach minimizes this problem because it strives to exhaust the literature through iterative searches, that is to collect information on the entire population. In our own study, we feel confident that the final set of coded publications includes virtually the entire population of publications. This may be concluded from the pattern shown in Figures 1 and 2.

Quality, Not Just Quantity

All meta-analyses are both work intensive and time-consuming (Berman and Parker 2002). In light of this, one might ask whether the sources uncovered in the latter search waves justify the additional effort.⁸ In other words, are these latter sources of poor quality? Quality assessments of the search waves show this is not the case. We assessed publication quality using three distinct measures. First, the coders assigned a three-level subjective rating to each publication in both the stress and social relationships projects. Publications were rated as low quality if they contained obvious reporting errors or applied statistical methods incorrectly. Publications were rated as high quality if models were well-specified – the correct model was used relative to the state of the art at the time of publication – and discussions and reporting of study results were detailed. Second,

principal components factor analysis was used to construct a 10-point scale quality measure for each individual meta-analysis (widowhood and mortality, group participation and mortality). Third, quality was assessed on the unemployment meta-analysis from the stress project using the Newcastle-Ottawa scale for nonrandomized trials (Wells et al., 2009). For this quality measure, the two coders independently rated each publication, using the average from their two ratings.

Subjective Quality. For the stress project, the mean subjective quality rating for all 262 publications was 2.37. A comparison of the different search waves shows that publication quality was actually higher in the later search waves rather than lower. The mean subjective quality rating was 2.30, 2.27, 2.23, 2.53, 2.43, 2.58, 2.57, and 2.78 respectively for each of the eight waves. The same picture emerged for the social relationships project. The mean subjective quality rating for all 329 publications was 2.59. Again, examinations by waves show that quality increases rather than decreases (the mean subjective quality was 2.15, 2.52, 2.64, 2.62, and 2.90 for each of the five respective waves).

Factor Analysis Quality Scale. Principal components factor analysis was used to construct the 10-point scale quality measure for each individual meta-analysis using the following: (1) the five-year impact factor (ISI Web of Knowledge, 2009) of the journal in which the article was published (an impact factor of 1 was assigned when the impact factor was not available); and (2) the number of citations received per year since publication, according to ISI Web of Knowledge. We present results from three individual meta-analyses to illustrate the relationship between search wave and scores on the 10-point quality scale. In our meta-analysis of group participation (church attendance, volunteering) and mortality ($n=93$), the mean quality score was 1.70. The mean quality score for each of the respective five search waves was 1.72, 1.61, 1.80, 1.46, and 2.02. As with the subjective quality rating, average publication quality did not decline in later waves.

Although the overall conclusion that the measure-driven approach yields high quality articles throughout the later search waves is supported, we obtained slightly different results with two other meta-analyses. In our study of war stress, the mean quality score among all 29 publications was 2.47. The scores increased over the first four search waves (1.08, 1.49, 2.90, and 3.47 respectively), but then declined over the remaining three waves (2.09, 1.48, and 1.83). In our study of unemployment, the mean quality score among all 42 unemployment publications was 2.85. The mean quality score was relatively stable over the first six search waves (3.53, 3.06, 2.89, 1.95, 2.37, and 3.68), declining only in the last two waves (1.54 and 1.32). It is important to note that in both meta-analyses, the last search waves, for which the quality score was lower, produced very few publications ($n=4$ for the last three waves of the war stress analysis; $n=4$ for the last two waves of the unemployment analysis). In both analyses, the bulk of publications were found in the middle search waves, where quality was at its peak. Without using the measure-driven approach, we would not have captured many of these high-quality publications.

Newcastle-Ottawa Scale (range, 0-9). We compared the mean scores of publications in each of the search waves of the unemployment analysis. The mean among all 42

publications was 7.51. There was exceptionally little variation in this quality score across the first seven search waves (7.57, 7.71, 7.75, 7.20, 7.00, 7.88, and 7.50). In the last search wave, which produced only one publication, the quality score was 4.00. Once again, these results show that the later search waves produce high-quality publications.

Quantity Matters – Exploring New Theoretical Dimensions Using Data Heterogeneity

Given the high time cost of the measure-driven approach, one may still doubt its substantive value. The addition of more quality data does not in itself guarantee that the results will substantively differ from those produced by smaller sample coming from a greater reliance on a concept-driven approach. However, meta-analyses of marital status and mortality (important subsets of our two projects) show that the addition of publications identified in later search waves provides important analytic leverage. Below we compare the findings from our own meta-analyses to those of Manzoli et al. (2007), who conducted a recent meta-analysis of marital status and mortality. We then show that the larger data set obtained through the measure-driven approach allowed us to explore moderating factors and other potential sources of data heterogeneity that Manzoli et al. could not examine.

To ensure comparability with the meta-analysis conducted by Manzoli et al., we replicated their selection criteria: We selected only studies of the elderly (mean age ≥ 65) published between 1994 and 2005 and adjusted for age, sex, and additional covariates. We found mean hazard ratios of 1.16 (95% CI, 1.09-1.23) for widowed persons, 1.14 (95% CI, 1.09-1.20) for divorced/separated persons, and 1.04 (95% CI, 0.99-1.11) for the never married. These mean hazard ratios are quite close to those reported by Manzoli et al., who found a mean relative risk of 1.11 (95% CI, 1.08-1.14) for the widowed, 1.16 (95% CI, 1.09-1.23) for the divorced/separated, and 1.11 (95% CI, 1.07-1.15) for the never married. In comparison to the results of Manzoli et al., our estimates are more conservative for the never married group, showing no significant effect among this elderly group.

As shown by the above comparisons, larger numbers of studies do not guarantee substantial differences in the mean effect being studied. That being said, the measure-driven approach almost invariably increases researchers' ability to explore sources of data heterogeneity, as demonstrated by our other findings from the marital status analyses. The ability to retrieve a much larger number of publications enables more sensitive analyses (using meta-regressions and sub-group meta-analyses) of important moderating factors such as age, gender, social class, and cultural characteristics, as well as analyses of an array of study design characteristics such as the date of data collection, composition of the groups being compared, etc. For example, in our meta-analysis of widowhood and mortality (Shor et al., 2012a), our sub-group meta-analyses and our meta-regression analysis detected a significant trend in the level of excess mortality risk following widowhood. The amount of excess risk was significantly higher in the first two years of follow-up, stabilizing at a lower level for the third year and on. Additional examples of findings like this one from our meta-analyses of marital dissolution (Shor et al., 2012b), lifetime non-marriage (Roelfs et al., 2011a), unemployment (Roelfs et al.,

2011b), religious participation (Shor and Roelfs, in press), war stress (Roelfs et al., 2010), and more are available from the authors.

Furthermore, the ability to retrieve a much larger number of publications facilitates the examination of interaction effects (a benefit of using the measure-driven approach which is rarely attempted in standard meta-analyses). The widowhood meta-analysis mentioned above provides an example for this advantage as well. Figure 7 shows an interaction of gender and mean age. These results show that, at younger ages, the excess mortality risk associated with widowhood is substantially greater for men than for women. However, the results also show that this excess risk declines more rapidly with age for men than it does for women. Such level of sensitivity provides important insights and substantially expands our existing knowledge of complex social phenomena. Individual studies and meta-analyses that have a smaller scope often fail to uncover patterns such as these.

Conclusion

We began this article showing that in recent years the use of meta-analysis in sociology is rapidly growing. This fact entails paying more attention to meta-analytic search and analysis techniques, ensuring that they are well adapted to the discipline. Conventional procedures of conducting meta-analysis often follow what we have called a concept-driven approach. While this approach is well-established and often useful, we showed that it suffers from a number of important shortcomings, especially when used in the social sciences. These shortcomings, we argued, are addressed by a more measure-driven approach, which we explicated here. Using two large projects, we demonstrated the strengths of this novel approach for meta-analytic sociological research. First, it forces the researchers to continuously revisit their conceptual vocabulary. Second, it substantially increases the number of sources found without compromising publication quality, thereby minimizing selection bias and allowing the researcher to further explore the sources of data heterogeneity.

The growing use of meta-analysis in sociology, along with the fragmented nature of the discipline, calls for increasing reflexivity regarding common methodologies and their appropriateness for sociologists. As Burawoy (1998) and others have argued, maintaining an open dialogue between theory and data is crucial for sociological research.⁹ In this article, we demonstrated the relevance of these insights for sociological meta-analytic research, both conceptually and practically. The measure-driven approach described here is an important step toward achieving a better balance between theory and data. While many meta-analysis researchers in the social sciences already recognize the advantages of such an approach, it is important for research reports and methodology publications to explicitly report and advocate the use of these conceptualizations and search methods.

In practice, sociological researchers will seldom initiate a search process that is strictly concept-driven or measure-driven. However, it is vital that the choice along the continuum of approaches be made explicitly and in a manner appropriate to the research terrain. Given the relative lack of conceptual unity and the frequency with which a given variable is used to represent multiple distinct theoretical constructs, we argue that

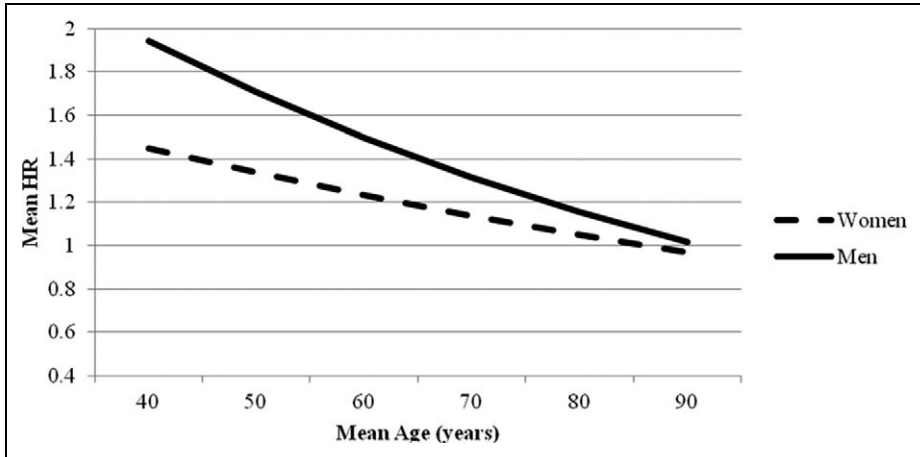


Figure 7. Mean HR by Mean Age and Gender.

sociological researchers are likely to benefit from pursuing an approach that is closer to the measure-driven end of the continuum.

Funding

The authors are grateful for the support provided by grant HL-76857 from the National Institutes of Health.

Notes

1. The terms “systematic review” and “meta-analysis” are often misused in the social sciences and elsewhere, being treated as synonymous with “thorough literature review”. When we refer to “meta-analysis”, a concept first introduced by Glass (1976 and 1977), we mean the strictly quantitative procedure that produces a weighted effect estimate (and usually a standard error or confidence interval).
2. For a more thorough presentation of these two approaches and the literature addressing meta-analysis search techniques, please consult the Online Supplement to this paper, available to *BMS* subscribers and RC33 members by distribution over the BMS-RC33 listserv at <https://listes.services.cnrs.fr/www/admin/bms-rc33>.
3. No previous study has provided an explicit, comprehensive definition of psychosocial stress. We therefore began with a working definition, which was later revised/refined based on an extensive survey of independent variables routinely categorized as psychosocial in the literature.
4. Studies which confounded all-cause mortality with one or more other health outcomes were excluded.
5. The numbers presented for the concept-driven search approach were derived in the following way: In the stress project, we had 30 articles coded from the original search and 16 articles coded from the subsequent bibliography search. In the social relationships project, the numbers were 14 and 40 respectively.

6. When restrictions on publication date are not considered, the number of articles found is much larger.
7. Of the 120 publications found, 40 had no more than 2 keywords of interest.
8. We estimate that choosing the measure-driven approach added about one additional year. The search and coding process took about 2.5 years at the rate of 10 person-hours per week. We estimate that using the traditional search procedure would have taken about 1.5 years.
9. While this article was written with the sociological researcher in mind, we believe that the measure-driven approach developed here will also be of use to researchers conducting meta-analyses in other disciplines. Meta-analysis is already heavily utilized by demographers, epidemiologists, public health scholars, economists, and many others. Our case studies overlapped with some of these fields and show that the literature searches conducted in these disciplines could also be expanded.

References

- Bailar JC (1997) The Promise and Problems of Meta-Analysis. *New England Journal of Medicine* 337(8): 559-61.
- Becker HA and Sanders K (2006) Innovations in Meta-Analysis and Social Impact Analysis Relevant for Tech Mining. *Technological Forecasting and Social Change* 73(8): 966-80.
- Berman NG and Parker RA (2002) Meta-Analysis: Neither Quick nor Easy. *BMC Medical Research Methodology* 2: 10-18.
- Burawoy M (1998) The Extended Case Method. *Sociological Theory* 16: 4-33.
- Egger M and Davey-Smith G (1998) Meta-analysis: Bias in Location and Selection of Studies. *British Medical Journal* 316(7124): 61-66.
- Farrington DP and Petrosino A (2001) The Campbell collaboration Crime and Justice Group. *The Annals of the American Academy of Political and Social Science* 578(1): 35-49.
- Glaser BG and Strauss AL (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Publishing Company.
- Glass GV (1976) Primary, Secondary and Meta-analysis of Research. *Educational Research* 5: 3-8.
- Glass GV (1977) Integrating Findings: The Meta-analysis of Research. *Review of Research in Education* 5: 351-79.
- Greenhalgh T and Peacock R (2005) Effectiveness and Efficiency of Search Methods in Systematic Reviews of Complex Evidence: Audit of Primary Sources. *British Medical Journal* 331: 1064-65.
- Higgins JPT and Green (eds) (2008) *Cochrane Handbook for Systematic Reviews of Interventions, Version 5.0.1*. New York: John Wiley, Cochrane Collaboration series.
- Huedo-Medina TB, Sanchez-Meca J and Marin-Martinez F (2006) Assessing Heterogeneity in Meta-analysis: Q Statistic or I2 Index? *Psychological Methods* 11(2): 193-206.
- ISI Web of Knowledge (2009) Journal Citation Reports. At http://wokinfo.com/products_tools/analytical/jcr/.
- Lipsey MW and Wilson DB (2001) *Practical Meta-Analysis*. Thousand Oaks, CA: Sage.
- Manzoli L, Villari P, Pirone GM and Boccia A (2007) Marital Status and Mortality in the Elderly: A Systematic Review and Meta-Analysis. *Social Science and Medicine* 64(1): 77-94.

- McCullough ME, Hoyt WT, Larson DB, Koenig HG and Thoresen C (2000) Religious Involvement and Mortality: A Meta-Analytic Review. *Health Psychology* 19(3): 211-22.
- Moher D, Liberati A, Tezloff J and Altman DG (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine* 6(7): 1-6.
- Roelfs DJ, Shor E, Davidson KW and Schwartz JE (2010) War-related Stress Exposure and Mortality: A Meta-analysis. *International Journal of Epidemiology* 39: 1499-509.
- Roelfs DJ, Shor E, Kalish R and Yogev T (2011a) The Rising Relative Risk of Mortality among Singles: Meta-analysis and Meta-regression. *American Journal of Epidemiology* 174(4): 379-89.
- Roelfs DJ, Shor E, Davidson KW and Schwartz JE (2011b) Losing Life and Livelihood: A Systematic Review and Meta-Analysis of Unemployment and All-Cause Mortality. *Social Science & Medicine* 72(6): 840-54.
- Shor E, Roelfs DJ, Curreli M, Clemow L, Burg MM and Schwartz JE (2012a) Widowhood and Mortality: A Meta-analysis and Meta-regression. *Demography* 49: -606.
- Shor E, Roelfs DJ, Bugyi P and Schwartz JE (2012b) Meta-analysis of Marital Dissolution and Mortality: Reevaluating the Intersection of Gender and Age. *Social Science and Medicine* 75(1): 46-59.
- Shor E and Roelfs DJ (In Press) The Longevity Effects of Religious and Non-religious Participation: A Meta-analysis and Meta-regression. *Journal for the Scientific Study of Religion*.
- Strauss AL and Corbin J (1990) Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Newbury Park, CA: Sage.
- Strauss AL and Corbin J (1994) Grounded Theory Methodology: An Overview. In Denzin NK and Lincoln YS (eds) *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.
- Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M and Tugwell P (2009) The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in Meta-analysis. At http://www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf.