

Reviews in Research on Adolescence: Genres, Trends, and Challenges

Roger J. R. Levesque¹

Received: 11 December 2015 / Accepted: 12 December 2015 / Published online: 21 December 2015
© Springer International Publishing 2015

Abstract The past few decades have witnessed rapid growth in research centered on the period of adolescence. Several markers of that growth are now obvious, such as multidisciplinary journals dedicated to that developmental period. But, it remains to be determined, like other fields of research that often contribute to the understanding of adolescence (genetics, cognitive science, medicine), whether multi-disciplinary developmental research on adolescence systematically evaluates the integrity of its findings. This study explores the extent to which the leading “adolescent” journals evaluate the state of knowledge that they disseminate, the importance of such evaluations, and challenges to the discipline’s scientific integrity. The review reveals that developmental journals devoted to the adolescent period pervasively do not publish articles that focus on reviews of research; it also finds that the most dominant forms of reviews are those deemed less rigorous. The discussion focuses on these findings’ implications and underscores the legitimacy and importance of doing what more established disciplines recognize as important and do more routinely.

Keywords Adolescence · Literature reviews · Replication · Reviews

Introduction

A dramatic growth in research outlets focusing on the adolescent period has emerged over the past few decades, and especially the last few years. The field now associated with the

study of adolescence as an academic, multidisciplinary discipline emerged during the 1960s and 1970s. It began with the founding of three multidisciplinary journals: *Youth and Society* (1969), *Journal of Youth and Adolescence* (1972), and the *Journal of Adolescence* (1978). In the 1980s and 1990s, three other important journals emerged, namely the *Journal of Early Adolescence* (1981), *Journal of Adolescent Research* (1986), and *Journal of Research on Adolescence* (1991). The field then stabilized, as the journals increased in volume rather than in number. In the new millennium, most of these journals increased the number of articles that they published. For example, the *Journal of Youth and Adolescence* now publishes over 190 manuscripts per year, a sizable increase compared to 40 in the year 2000, and 57 in 2005; similarly, the *Journal of Adolescence* has more than doubled in volume, from 63 articles in 2005 to over 150 in 2014. Without doubt, these journals exhibit and support a rapidly expanding field.

The expansion in the number and volume of publications has been paralleled by developments in other disciplines that also feature research on adolescence. Several disciplines now support journals that focus on different aspects of adolescence, such as adolescent health in public health, pediatrics in medicine, juvenile justice and delinquency in criminology, developmental science in psychology, and childhood in education, sociology, neuroscience and human development. For example, the Society for Research on Child Development supports multiple journals that group adolescents with children. Larger professional organizations, like the American Psychological Association, American Medical Association, American Educational Research Association, and American Sociological Association, also have journals focusing on youth as well as on contexts in which they develop (e.g., schools, families, and health systems). In any given year, these outlets produce literally thousands of empirical studies relating to the adolescent

✉ Roger J. R. Levesque
rlevesqu@indiana.edu

¹ Indiana University, 302 Sycamore Hall, Bloomington, IN 47405, USA

period. As a result, the diffusion of research on adolescence is broad, deep, immense, and briskly expanding.

The status of adolescent research reveals a vibrant field but the dispersion of research findings makes unclear what we actually know about particular topics. What is known and not known remains obscured without authoritative statements that cohesively bring together key findings. The obscurity is especially pronounced given the field's multidisciplinary focus, which means that research often comes from different disciplines that can have very different publication standards, embrace different methods, and focus on different aspects of phenomena.

Despite the above developments, what exactly makes for authoritative reviews remains a matter of dispute. Even what constitutes a review remains debatable. Although many fields of empirical inquiry have recognized these issues, that is not the case for research centering on adolescence. The field of adolescent research has not addressed how its reviews can leave much to be desired and it even leaves unclear what would be desired.

This study addresses the failure to examine the nature of reviews focusing on adolescent development. After highlighting the importance of rigorous reviews, the article explores the general genres of reviews. Then, it examines trends in the use of identified genres in the leading journals dedicated to the multidisciplinary understanding of adolescence. As it does so, the discussion reveals the common and ignored forms of reviews, gauges the relative quality of these reviews, and identifies key challenges. The discussion concludes by revisiting lessons learned and by identifying steps forward.

The Changing Importance of Literature Reviews

A discipline cannot grow effectively without a firm foundation of prior knowledge. The development of empirical fields of study relies on the accumulation of knowledge that maintains and improves its scientific rigor. Effective programs of research must look to prior knowledge to understand phenomena. When doing so, researchers must synthesize the empirical evidence, develop theories, provide conceptual backgrounds for their research, and identify topics or research domains that require more investigation. In essence, researchers must take research findings and integrate them into research programs.

The integration, theory building, and hypothesis testing expected from rigorous research always have benefited from reviews. But, research programs and disciplines now benefit even more from compelling reviews that either stand on their own or serve to launch other investigations. Reviews benefit research because of changing trends in scientific publications and concerns about the dissemination of accurate scientific findings.

The importance of reviews partly comes from the increasing move toward shorter empirical manuscripts and increased demands to provide readers with more details about their methods and analyses. Detailed foundations for hypotheses or research questions are being replaced by short statements sending readers to other research, which often is research that adopted similar approaches to revealing the foundation for their own studies. These trends mean that researchers and others interested in knowing a topic's state of the art increasingly must rely on reviews that bring together disparate studies into authoritative statements. The need for such statements now has been recognized by many disciplines—the large disciplines supported by professional associations listed above—with journals that review the literature as well as journals that examine how to review empirical literature.

In addition to trends in publishing research, the need for reviews comes from the discipline's need to evaluate its findings. Researchers have become increasingly concerned about revelations that highly regarded studies have failed to meet scientific expectations of reproducibility. For example, a recent article in the highly regarded scientific journal *Nature* reported that Amgen, a major American biopharmaceutical company, sought to replicate the findings of 53 published cancer research studies that it deemed “high profile;” it was able to reproduce only 11 % (6 of the 53) (Begley and Ellis 2012). Similar findings led the US National Institutes of Health (NIH) to undertake initiatives to enhance reproducibility (Collins and Tabak 2014) and have led to large-scale efforts to replicate studies from leading scientific journals (see, e.g., Open Science Collaboration 2015). Although heightened, these concerns have been discussed among scientists for years, as scientific disciplines always have been concerned about the necessity to evidence replication (Campbell and Jackson 1979). The concerns go to the core of the scientific method: Outcomes of research, although statistically significant, are not factual unless they can be replicated. Without replication, which serves substantiation and verification functions that increase validity, it is difficult to place credence on one perspective or another.

Despite the need for replication, inherent problems stymie efforts to conduct them. Commentators, for example, report an unwillingness or inability to share published data, fewer published replications than in the past, and questionable research practices (Pashler and Wagenmakers 2012). They also reveal that journal editors, particularly those in the social sciences, do not evince much enthusiasm for replications (Easley et al. 2013), a finding supported by reviews of disciplines that identify very few published replications (e.g., the complete publication history of the current top 100 education journals ranked by 5-year impact factor found that only 0.13 % of education articles were replications; see Makel and Plucker 2014). Yet, such

findings should not be surprising as academics place a premium on originality. Some of the driving forces behind many publications—tenure, promotion, professional recognition—rest on innovative work. The actual practice of the field, then, suggests that novelty appears more important than scientifically acceptable facts.

The challenges of conducting and reporting replications make the need for careful reviews of the literature even more important. In the absence of direct replications, researchers rely on (1) statistical methods to control rates of false positives, (2) conceptual replications, and (3) the self-correcting nature of science. Each of these approaches is problematic (see Pashler and Harris 2012; Shuster and Cottrill 2015), and some commentators conclude that direct replication is the only way to verify the reliability of findings (Simons 2014). Despite scientific ideals of replication, however, some methods are more feasible in practice, such as now well-established alternatives to direct replication. Some commentators even argue that these alternate methods are superior, as they champion conceptual replications that focus on testing theoretical hypotheses rather than assessing the reliability of particular experimental procedures (Stroebe and Strack 2014). Not surprisingly, others concerned about replicability issues conclude that “research on research” (i.e., meta-research) now serves as a key way to determine the inadequacies of findings (Ioannidis 2014, p. 3), which is a reasonable argument given that a meta-analysis actually is an explicit form of external replication. It also is reasonable because even replications need replications—a replication study that finds results different from an original study does not necessarily negate the original findings. Such observations, particularly in light of the use of other corrective measures of individual studies (see below), reveal the important role high-quality reviews of existing research plays in shaping scientific disciplines.

The Nature and Quality of Literature Reviews

The most recent and comprehensive examination of reviews of what constitutes literature reviews (see, e.g., Paré et al. 2015) reveals a wide range of definitions and descriptions. These differences have led to calls for greater scrutiny of the methods used to review and for efforts to set standards to distinguish between inferior and superior reviews. Yet, these efforts generally have not gained much traction. The failure to gain traction results from many factors, but it likely mainly comes from the nature of available research to review, the methods available to analyze primary research, and the needs of those attempting to make sense of available studies. Despite these limitations, the reviews themselves become the sign of the relative vigor of fields of study as they bring together useful research that sets the stage for more.

One notable exception stands out from the general inability to determine the relative rigor of reviews. That exception involves reviews that integrate and aggregate empirical findings. This exception emerged from the introduction of meta-analytic procedures that now dominate writings of research syntheses and even journals devoted to synthesizing research (see *Research Synthesis Methods*, founded in 2010). Yet, these types of systematic reviews actually comprise only a very small percentage of published reviews. Established research programs based on multiple and/or large data sets are best suited to benefit from the currently developed empirical tools. As a result, the tools championed as the gold standard for reviews remain unsuitable for the vast majority of reviews that do not draw from already fine-tuned analyses and established research programs.

Given the wide variety of studies that can be reviewed, and the lack of standards to mark their effectiveness, researchers who study the nature of reviews have sought to describe and categorize them into different types. They have done so with the hope that detailing the nature of reviews could increase the effectiveness of reviews and assist in evaluating their relative effectiveness. Those prior efforts (as shown below) reveal that they tend to describe and understand reviews in terms of their sources, goals, and methods of evaluation.

Approaching reviews by focusing on their sources, goals, and methods appears straightforward enough, but doing so has important limitations. Such approaches can mask many layers of complexity, nuances, and even contradictions. Generally, for example, reviews’ sources are primary or original scholarship; reviews do not report primary scholarship. Yet, that general rule has many exceptions. Reviews can create new scholarship depending on the methods that they use to report. For example, they can aggregate statistical results of prior research, which would provide new primary data; and they can tabulate trends, which would provide original data. In addition to variation in what reviews produce, what constitutes primary research can range widely. Primary research need not involve empirical data, as it can include theoretical, interpretive, or conceptual analyses as well as methodological analyses. Goals also can vary tremendously, and particular reviews could embrace multiple goals at once. For example, reviews can seek to describe, evaluate, clarify, summarize, and/or integrate. Lastly, methodologically, much depends on the sources of data and the goals for the reviews; and sometimes the reviews may adopt multiple methods depending on sources and goals. In the end, meeting a particular goal or using a particular method may not reveal much about the authoritativeness of the review. Reviewing reviews involves making judgment calls, an approach that lends itself to criticism but remains needed to move the field forward.

Researchers and commentators, then, may have a clear sense of what constitutes high quality reviews, but the

quality of the vast majority of reviews remains difficult to discern. The focus has been on fine-tuning the most rigorous form of reviews—systematic quantitative reviews—rather than on understanding the value of other approaches and enhancing their quality. This leaves the next step in understanding and improving the wide variety of reviews as beginning by detailing the possible range of reviews, including how they rely on original sources, what they seek to do with them, and what they do with them. Focusing on these factors still has limitations, but it provides a reasonable starting point that can provide a sense of the reviews' overall quality.

Genres of Literature Reviews in Research on Adolescence

Reviews can be categorized in many ways. One recent and thorough review of reviews, for example, identified 40 sources that reported different categories of reviews. To do so, it first described what it considered articles that articulated the nature of reviews. The reviews of reviews would need to (1) identify, describe and often apply different types of reviews; (2) make original contributions to the theory of literature reviews; or (3) be particularly influential (see Paré et al. 2015). This approach resulted in a broad standard that casted a broad net. Although the search was broad, its reference section identified no sources from developmental research. Efforts to categorize and describe reviews mainly came from health, education, management and information sciences.

Although not identifying work from developmental science, the sources from the reviews provided useful analyses. First, they distinguished nine different types of reviews. Those nine types are common across other analyses of types of reviews. Second, the nine types are manageable in that they permit a reasonable amount of nuances and differentiation. That differentiation works better, for example, than typologies that identify large numbers of categories, since they would not work well in fields with potentially small numbers of reviews (e.g., fourteen; see Grant and Booth 2009). Third, the sources indicate how different fields develop or even prefer different types of reviews. Lastly, reviews of reviews essentially reveal an immense diversity across fields of study. The fundamental lesson that emerges is that the diversity of reviews across fields of research makes it inadvisable to simply transplant onto one field another field's approach.

Despite rigorous efforts to develop typologies, existing ones remain unsatisfactory for at least two reasons beyond the diversity, number, and particular need for those that do exist. First, although conceptually tidy, they are difficult to use in that published studies often exhibit multiple aspects

of review types, such as a variety of data synthesis methods, which makes it difficult to categorize the reviews. Essentially, existing typologies have reliability problems. Second, despite their exhaustiveness, even reviews of typologies fail to identify important types of reviews published in journals focusing on the period of adolescence. Even a passing familiarity with journals focusing on the period of adolescence reveals, for example, peer commentaries and book reviews that tend to go ignored by typologies of methods for synthesizing research. Thus, existing typologies also have validity problems in that they ignore much of what they say they are measuring, in addition to lacking validity in terms of not transferring well from one field to another.

The diversity and limitations of existing typologies provide the impetus for taking a close look at how the field of adolescent research approaches reviews. Available typologies, coupled with openness to identifying new forms of reviews, serves as a reasonable starting point to identify trends, gaps, and potential areas ripe for development. In fact, typologies have identified this approach as a reasonable one and even have created a category for it, as seen below.

Using the above approach, articles in journals focusing on the adolescent period were scanned to determine their fit with different types of reviews. That preliminary review confirmed that typologies described in the literature on reviews do not transfer well to research on adolescence. Only three of the typical types of reviews appear to transfer, and those individual types need to be expanded to include more diverse forms of reviews. And, as already noted, the journals publish types of reviews that have not even been recognized as reviews. The upshot is that the nature of the data simply does not support the use of several clearly demarked categorizations identified by reviewers of reviews. It may well be that useful typologies could be developed later, but the current state of the field, including how reviews are conducted, tends to support broad categories akin to more fluid genres. This section describes those genres (see Table 1) and then distinguishes several other forms of reviews that have yet to gain much traction. It does so to get a sense of the field and its potential.

Genres of Reviews in the Study of Adolescence

Narrative Reviews likely are the most readily understood, as they are the most commonly reported form of review identified in studies reviewing reviews. They are similar to reviews of the literature that empiricists use to build and test their hypotheses. Narrative reviews can vary tremendously in form but, at their core, they identify what has been written on a subject. Typically, these reviews tend to be opportunistic and use evidence available to the researcher, do not provide information as to how the studies were identified,

Table 1 The typical nature of review genres in adolescent research

Review genre	Goal	Examples	Search strategy	Explicit study selection	Quality appraisal of original studies	Methods of analysis
Narrative Review	Summation of prior knowledge	Narrative	Typically selective	No	No	Narrative summary
		Descriptive	Representative	Yes or no	No	Content or frequency analysis
Systematic Review	Data aggregation and integration	Meta-analysis	Comprehensive	Yes	Yes	Statistical, meta-analysis
		Qualitative synthesis	Comprehensive	Yes	Yes	Narrative synthesis
Critical/Realist	Explanation building	Theoretical	Selective or representative	Yes or no	No	Content or interpretive analysis
		Conceptual	Selective or representative	Yes or no	No	Content or interpretive analysis
Peer Commentary	Evaluation	Commentaries	Typically none	No	Yes or no	Narrative analysis
Book Review	Summation and evaluation	Book Reviews	Typically none	No	Yes or no	Narrative analysis

This table benefitted from many prior analyses, particularly Paré et al. (2015)

and are selective rather than systematic and comprehensive. They seek to use compelling examples to support their points. As a result, narrative reviews can be very much like commentaries and other methods of reviews that develop subjective interpretations of available evidence. By using informal techniques that are neither revealed nor justified, narrative reviews leave readers to infer credibility by the extent to which they find the narratives compelling. The lack of explicit methods and reliance on subjective approaches to synthesizing research creates important limitations. Like other methods, they also can gain credibility to the extent that they leave readers with a sense that the method taken best fits the available data.

Systematic Reviews now likely are what come to mind when thinking of compelling literature reviews. Researchers have identified many types of systematic reviews, but *Meta-analytic Reviews* and *Qualitative Systematic Reviews* are the most common. Meta-analytic reviews make use of statistical methods to contrast and combine results from different studies in order to identify patterns among study results, sources of disagreement among those results, or other relationships that may come to light in the context of multiple studies. These reviews aggregate data findings to increase the statistical power for the measure of interest, which can reveal significant findings from pooled data that otherwise would not have been significant. Meta-analyses are seen as powerful tools that can be transparent, for example, in the manner that they set criteria used to search for studies, select studies based on objective criteria, deal with incomplete data, and account for or ignore publication bias. The transparency is

helpful, but it still leaves important discretion to analysts. That concern for discretion appears even more pronounced with qualitative systematic reviews, which have as their defining characteristics textual approaches to analyses rather than quantitative analyses. Instead, they use classification schemes, frameworks, and other ways to group findings and present compelling analyses. Like meta-analyses, systematic qualitative reviews seek to be transparent in terms of the parameters for their searches and analyses. Still, they retain much subjectivity in the narrative and textual analyses that they support.

Critical/Realist Reviews constitute the third major form of review. Sometimes typologies separate these types into two different ones, while at other times they place them together and do not identify much difference. When applied to research on adolescence, it makes sense to place these approaches together for the simple reason that it is difficult to distinguish between both and very few reviews fall under the category. These reviews seek to explain phenomena by articulating likely underlying mechanisms and then analyzing available evidence (qualitative, quantitative, or theoretical) to determine whether and when the mechanisms are applicable. The method typically relies on preliminary literature searches that map out issues that are then analyzed in greater depth to highlight the relative support for claims. Such reviews can find parts of flawed studies helpful, as the goal is to identify mechanisms that explain phenomena and the relative support for the mechanisms. Unlike systematic reviews, these reviews need not reveal the methods used for synthesizing evidence, the

standards used to appraise the rigor of the primary studies, and the identification of the original studies. Sometimes typologies of reviews distinguish critical from realist reviews by highlighting that critical reviews focus on identifying weaknesses, contradictions, inconsistencies, and controversies. By seeking to highlight problems, they do not exhaustively review existing studies, rather they select particular studies or identify representative ones. They also use studies that adopted different analytical methods. Just as with realist and narrative reviews, they are vulnerable to charges of subjectivity.

A close analysis of the content of journals focusing on adolescence reveals that existing typologies ignore an important form of reviews: *Peer Commentaries*. Commentaries vary considerably in approach, but they clearly are reviews. They can be distinguished from other types of reviews by the manner that they respond to specific works, be they articles or even programs of research that take sides on controversies among experts. Typically, the reviews are not exhaustive, as they focus on the presented work to which they are responding. Peer commentaries also are evaluative and similar to critical reviews in that they point to weaknesses and strengths in other people's work. Peer commentaries sometimes can be distinguished from other reviews in that they can present alternative ways to understand or view the findings under scrutiny.

Perhaps because fields vary in the extent to which they value articles and books as venues for presenting research, existing typologies ignore another common form of reviews: *Book Reviews*. These reviews also can vary in how they develop and present their reviews. Sometimes they can be quite short and offer a synopsis of the book. At other times, they can be more thorough and describe books' contents, either in summary form or in an evaluative manner. At other times, books serve as springboards for authors to develop and present alternative views. Sometimes books are reviewed alone, while at other times several are grouped together. As a result, this method of review can adopt analytical approaches similar to other genres of reviews, and it can even be quite similar to peer commentaries as they do just that: comment on peers' books. But, this genre of review remains generally distinguishable by its focus on books and, equally helpful to distinguish them, journals typically identify them as book reviews. In a real sense, this genre tends to be the most readily distinguishable, despite its being the most ignored by efforts to identify and document the nature of literature reviews.

Ignored or Insufficiently Distinguishable Review Genres in the Study of Adolescence

Consistent with some forms of reviews that highlight gaps, it is important to note some review types identified by prior

reviews but that have yet to be published in journals focusing on the adolescent period or that are not distinguishable enough from other genres of reviews. Four are important to note.

Among the most striking omissions are *Scoping Reviews*. Such reviews conduct broad searches to determine the size and nature of research on specific topics. They seek to be comprehensive and are not concerned with the quality of research, an approach that some view as limiting. Still, at some level, they actually appear to be a form of systematic review since they clearly state inclusion and exclusion criteria and systematically search the literature.

Descriptive Reviews would seem to be prominent, but they are not. They are said to represent the state of the art in a specific research domain. Using that standard, many reviews in adolescent journals would fall under this category. But, this type of review tends to be described as using numeric data to reflect the frequency of topics in the extant field. They also highlight the nature of the data and findings. For our purposes, those are insufficiently distinguishable from systematic reviews as they are, by their nature, systematic.

Umbrella Reviews also have been conducted in many fields, but not directly in adolescent research. This is not surprising, as umbrella reviews provide overviews of multiple reviews. Given the paucity of reviews found in adolescent journals, the state of knowledge appears to not have reached the stage that reviews can review multiple reviews.

Theoretical Reviews also do not map well onto what is being published in adolescent journals. They do not because they tend to develop conceptual models or frameworks. Doing so makes them very similar to other genres of reviews. Also, theoretical reviews can adopt multiple techniques to review materials. Despite the importance of theoretically guided research and theoretical applications, the field evinces little support for reviews evaluating theories.

The Current Study

Important lessons emerge from a preliminary look at the review genres appearing in multidisciplinary, developmental journals devoted to the study of adolescence. As can be seen from these commonly identified types of reviews, efforts to distinguish among review types eventually become unhelpful. Fine-tuning does not reveal much more about the research record than we already can determine by other genres. Still, it remains important to clearly delineate, as much as possible, the nature of the genres that are identified and to examine their utility to the field. Doing so highlights the challenges faced by efforts to create typologies. In the end, it makes more sense to use the term "genre," as genre can refer to categories of literature characterized by similarities in form, style, or subject

matter. Although creating a high number of clearly demarked typologies would seem to be warranted given that reviews seek to analyze empirical research, doing so simply would not be helpful.

To get a good sense of how the field is reviewing itself, this study applied the genres to the five journals broadly dedicated to the adolescent period: *Journal of Adolescent Research*, *Journal of Adolescence*, *Journal of Research on Adolescence*, *Journal of Youth and Adolescence*, and *Youth & Society*. As noted above, many types of journals do publish manuscripts relating to the adolescent period, but they are specialized, for example, toward specific fields (e.g., health, criminology, or education) or even specific periods of “adolescence” (e.g., early adolescence or emerging adulthood) or other groups (e.g., children). Assessing more journals would have been possible, but the goal was to get a sense of how the field approaches reviews. That goal led to a focus on the journals with the broadest inclusion of research on adolescence, both in terms of the age period (i.e., early adolescence to emerging adulthood) and disciplines (i.e., the journals are multidisciplinary).

This study sought to uncover more than the genres used in the field’s key publishing outlets. It sought to understand trends in the use of those genres. More fundamentally, it sought to understand the field’s response to concerns about scientific rigor. This response was deemed important to determine what the field views as publishable reviews and how to evaluate the effectiveness of reviews. It turns, these goals are important to shed light on how to shape the development of useful reviews as well as publications that will provide information necessary for effective reviews.

Methods

A pilot search in databases like ISI Web of Science, PubMed and PsycINFO revealed the inadequacy of relying on them to identify reviews, as many reviews do not identify themselves as such. As a result, the journals were manually searched. Before developing a coding scheme, the journals were quickly scanned to determine the scope of the search, in terms of journal volumes. Beginning with the new millennium made sense given that the journals were stable in 2000, the 15 period would cover a broad enough range to reduce the effects of editorial board changes, and it would provide a reasonable number of manuscripts to review (several thousand).

Using Table 1 and information about other types of reviews as guides, the journals were searched to develop a coding scheme. It immediately became clear that many types of reviews were not being used (such as scoping reviews). The search also revealed only a few meta-analyses, and that they could be categorized with other systematic

reviews. Other types of reviews, as identified above, were difficult to distinguish from each other (realist/critical and narrative/descriptive); and so they were coded together. The process left these reviews: *Narrative*, *Systematic*, *Critical/Realist*, *Peer Commentary*, and *Book Review*. Also notably, some manuscripts were not necessarily articles; some were, for example, brief one-page editorials. This led to not counting brief editorials as articles; brief reports of empirical findings, however, were counted as articles. They were mainly because the electronic presentation of journal issues for journals that do use brief reports count them as articles.

The manuscripts identified as reviews were then placed into their respective categories based on the particular review’s goals and method. Merging some of the more difficult to distinguish types together increased confidence in the coding scheme’s reliability. To ensure reliability, the analysis conducted by the author was redone by two research assistants. The assistants were given five copies of the different reviews and asked to identify them; their assessments were identical. The assistants then proceeded to work through each volume of the five journals to identify reviews and to identify the genre of the reviews; in a handful of instances where there was uncertainty about the coding, it was coded after discussing with the research team. The result was agreement on all of the coding.

The above efforts give us confidence in our coding scheme, but it is important to note that other research teams may end up with variations. Variations may be due to what they might count as an article. In addition, they also might count some of the genres differently because the research team needed consults for some manuscripts and made executive decisions; this was particularly true for systematic v. narrative reviews. Prior efforts simply have designated some of these situations as hybrid cases; but we opted to commit to one genre based on the team’s views of the particular review’s goals and their dominant method. Given the potential uncertainty, all narrative and systematic reviews were double-checked and coded by two team members, which was not an onerous task given their small number.

Results

The overall results of the search and coding are presented in Tables 2 and 3. Table 2 reveals considerable variety in what journals publish. For example, some publish no book reviews and some focus more on narrative and systematic reviews. Table 2 also reveals that the overall percentage of reviews is actually quite small, 338/4170, or 8%. Removing book reviews, which no existing typology counts as literature reviews, reveals an even smaller percentage: it cuts the original by half—151/4170, or 4

Table 2 Review genres by journals focused on adolescent research (2000–2014)

Journal	N	Sys	Critical	Peer	Book	No. of articles
JAR	10	0	2	2	22	480
JoA	3	11	5	13	63	1345
JRA	25	3	17	11	0	582
JYA	10	6	6	15	102	1369
Y&S	1	0	11	0	0	394
Total	49	20	41	41	187	4170

Journals: *Journal of Adolescent Research* (JAR), *Journal of Adolescence* (JoA), *Journal of Research on Adolescence* (JRA), *Journal of Youth and Adolescence* (JYA), and *Youth & Society* (Y&S). Review Genres: Narrative Review (N), Systematic Review (Sys), Critical/Realist Review (Critical), Peer Review (Peer), Book Review (Book)

percent. Table 2 also reveals that the least frequently published reviews are those deemed most rigorous—systematic reviews. Table 3 presents reviews over time. It reveals a general increase in the number of reviews, much of which could be attributed to book review essays.

The raw data, not presented in the tables, permits a few other notable observations. Journals published only a handful of systematic reviews before 2010. Similarly, very few narrative reviews were published before 2011, with the exception of the *Journal of Adolescent Research* that published most of its narrative reviews before 2011 and the *Journal of Adolescence* that published most of its narrative reviews before that time and, since then, appears to have switched to supporting systematic reviews. Also in terms of narrative reviews, the *Journal of Research on Adolescence*

published none before 2011 and the bulk of those that it did publish that year were “Decade in Review” pieces (some of those review pieces were deemed systematic). In terms of book reviews, the *Journal of Youth and Adolescence* published none until 2007, the *Journal of Adolescent Research* published theirs mainly from 2005 to 2011, the *Journal of Adolescence* published its book reviews from 2000 and stopped in 2009. Peer commentaries appear randomly but, when they do appear, they tend to appear in groups, such as commentaries on journal special issues. Lastly, *Youth and Society* publishes very few reviews; and those that were published tended to appear either 2001 or 2011. These findings reveal trends that are idiosyncratic to specific journals. There are no dramatic increases in any forms of reviews across journals.

Discussion

Several well-established fields of study have recognized the importance of reviewing their empirical findings. Before this study, it remained unclear whether, and if so how, the relatively young multidisciplinary study of adolescence was reviewing itself. This study sought to address this concern and, in the end, identified a pervasive lack of articles devoted to reviewing prior research. In doing so, the study underscores the need to address challenges that other fields have identified and create opportunities to better benefit from prior research.

The current state of reviews reveals much more than the discipline’s tendency to not review itself. It reveals limited

Table 3 Review genres in journals focused on adolescent research 2000–2014

Journal	N	Sys	Critical	Peer	Book	No. of reviews	No. of total articles
2014	3	3	6	10	21	43	439
2013	4	2	2	7	27	42	401
2012	5	4	–	1	16	26	419
2011	20	2	6	3	13	44	425
2010	2	2	2	5	12	23	315
2009	1	–	1	1	10	13	302
2008	3	–	–	4	17	24	249
2007	1	2	–	–	16	19	262
2006	–	2	2	1	9	14	230
2005	–	1	7	–	8	16	198
2004	1	2	–	–	8	11	204
2003	2	–	2	–	8	12	181
2002	2	–	5	1	10	18	183
2001	4	–	6	–	8	18	181
2000	1	–	2	8	4	15	181
Total	49	20	41	41	187	338	4170

Review Genres: Narrative Review (N), Systematic Review (Sys), Critical/Realist Review (Critical), Peer Review (Peer), Book Review (Book)

developments in evaluating the study of adolescence. One of the key findings is that the most prominent genres of reviews that are published are those that other fields do not view as robust. For example, a flagship journal in the discipline, the *Journal of Research on Adolescence*, published several “Decade in Review” articles that relied on narrative and less empirically rigorous forms of reviews. Those articles presumably benefitted from a variety of peer reviews and likely will impact the field. As such, the publication of these reviews highlights what constitutes acceptable reviews—reviews need not meet the gold standard in quantitative analyses to make contributions in their evaluation of the discipline. Although other journals have not published “Decade in Review” pieces, they too reflect the tendency to use less empirically rigorous review methods.

Research published in well-regarded journals necessarily use a wide variety of methods to accommodate the nature of materials reviewed and reasons for reviewing them. Yet, the diversity of methods used, especially of methods deemed less rigorous, still raises important concerns. First, reviews may not be well received by consumers of research. This lack of receptivity may be true for researchers as well as for those who would seek to enhance practice or shape policies requiring evidence-based research (see Nordmanna et al. 2012). Second, the reviews may provide inadequate conclusions and determining whether they are doing so would be impossible at least until researchers conducted more studies and reviews. This presents a significant challenge because it hampers the most routine aspects of evaluating and reporting empirical findings, such as peer review. The diversity of methods used without clear articulations of their effectiveness leaves the process open to biases and other problems that come from the difficulty of judging the quality of the reviews.

The challenge to ensure the appropriate development and publication of reviews has been faced by other disciplines and has resulted in calls to determine how best to evaluate reviews. Such calls typically invoke requests to increase transparency, e.g., reveal potential conflicts of interest and make approaches and intentions explicit (see, e.g., Cooper and Koenka 2012; McGaghie 2015). These are important fundamentals. They permit the use of flexible evaluations so that progress can be made in providing a sense of what we know about particular topics. But, as we have seen, the majority of reviews published in journals focused on the adolescent period currently do not consider these fundamentals (and, if they do, they are not discernible from the publications themselves). These inadequacies create difficult situations. It is one thing to call for flexibility, but it is an entirely other matter to call for low, no, or obscure standards.

The need to tell more than compelling stories highlights the most important challenge: concern about false findings. Recently, well established fields of study have sought to

address these issues not just in techniques used to review studies but also in the need for studies to, essentially, review themselves. For example, critical reviews of candidate gene and Gene \times Environment interaction (known as cGxE) studies published in the new millennium (e.g., Duncan and Keller 2011; Dick et al. 2015; Keller 2014) reveal that the majority of studies are underpowered, most likely significant with small sample sizes, rarely directly replicated and frequently indirectly replicated, fail to control for confounders (e.g., ethnicity, gender, age, socioeconomic status, etc.) rather than by the specified genetic or environmental variables per se, and likely result from publication bias whereby authors more likely submit (and editors more likely accept) statistically significant cG \times E findings. These reviews even have estimated that the majority of cGxE studies published since 2000 present false positives that do not stand up to subsequent replication attempts, a skepticism spurred by meta-analyses raising concerns about the quality of rapidly expanding genetic research (e.g., Risch et al. 2009).

Awareness of these limitations has led to suggestions to address them. At the forefront of efforts has been the study of behavior genetics. For example, the discipline’s leading disciplinary journals (Behavior Genetics; see Hewitt 2012) and journals that routinely publish genetic research (Journal of Abnormal Child Psychology; see Johnston et al. 2013) have published policies outlining stricter criteria that manuscripts reporting candidate gene main effects or interactions must meet before they will be considered for peer review. These criteria include the use of multiple statistical methods to evaluate an effect and following up significant cGxE effects with replication attempts in independent samples as well as replication within the same sample. Similar ways to address the potential limitations of genetic \times environment research have been championed by disciplines that may have more difficulty reaching the expected standards due to, for example, the nature of their datasets (see Salvatore and Dick 2015; Schlomer et al. 2015).

The many concerns raised by behavior genetics research are not new. They are similar to long-identified concerns about research integrity (i.e., high false-positive rates) that have received increasing attention in the social sciences (e.g., see analyses of cognitive science, Ioannidis et al. 2014; also see clinical science, Mueller et al. 2014; also see developmental science, Duncan et al. 2014). They also have been raised in journals dedicated to the adolescent period (Mullineaux and DiLalla 2015; Levesque 2015). However, prior responses to these types of challenges focused on, for example, making proper covariate adjustments, such as has been championed in personality (Hull et al. 1992) and social psychological (Yzerbyt et al. 2004) research. The new stricter criteria are much more challenging.

How disciplines will respond to mounting criticisms and challenges remains to be seen, but competent reviews (across studies and within them) necessarily will be part of effective responses. Criticisms reveal the need for careful reviews of research. They also highlight the benefits that can come from understanding the inadequacies of existing bodies of knowledge. That care and understanding is what moves calls for more transparency in primary research in order to increase the transparency and rigor of reviews.

Conclusion

The study of adolescence may have come of age, but one would not know it. Research remains widely dispersed across many disciplines. Several important journals dedicated to the adolescent period have yet to showcase programs of research in a manner similar to more established disciplines. A very small percentage of articles focuses on reviewing or otherwise evaluating multiple studies' findings. Reviews of the literature from other disciplines reveal that this lack of development in adolescent research limits the discipline's exposure, ability to develop, and potential influence. Concern rises as other disciplines reveal that the lack of direct/conceptual replications and reviews contributes to bad science.

Despite limitations in the manner the discipline reviews its findings, it is not possible to reach many definitive conclusions about the state of adolescent research given the limited developments in efforts to review and understand the discipline. Lack of knowledge means an inability to reach firm conclusions, which is precisely the reason that the discipline would benefit from paying more attention to the state of its own research. Building requires a strong foundation; in this instance, it requires knowing what is known and not known as well as what would be useful and why. Only then can the multidisciplinary study of adolescence address its potential limitations and ensure its integrity.

References

- Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531–533.
- Campbell, K., & Jackson, T. (1979). The role of and need for replication research in social psychology. *Replications in Social Psychology*, *1*, 2–14.
- Collins, F. S., & Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature*, *505*(7485), 612–613.
- Cooper, H., & Koenka, A. C. (2012). The overview of reviews: Unique challenges and opportunities when research syntheses are the principal elements of new integrative scholarship. *American Psychologist*, *67*, 446–462.
- Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., et al. (2015). Candidate gene–environment interaction research reflections and recommendations. *Perspectives on Psychological Science*, *10*, 37–59.
- Duncan, G. J., Engel, M., Claessens, A., & Dowsett, C. J. (2014). Replication and robustness in developmental research. *Developmental Psychology*, *50*, 2417–2425.
- Duncan, L., & Keller, M. C. (2011). A critical review of the first 10 years of measured gene-by-environment interaction research in psychiatry. *American Journal of Psychiatry*, *168*, 1041–1049.
- Easley, R. W., Madden, C. S., & Gray, V. (2013). A tale of two cultures: Revisiting journal editors' views of replication research. *Journal of Business Research*, *66*, 1457–1459.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*, 91–108.
- Hewitt, J. K. (2012). Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behavior Genetics*, *42*, 1–2.
- Hull, J. G., Tedlie, J. C., & Lehn, D. A. (1992). Moderator variables in personality research: The problem of controlling for plausible alternatives. *Personality and Social Psychology Bulletin*, *18*, 115–117.
- Ioannidis, J. P. A. (2014). How to make more published research true. *PLoS Med*, *11*, e1001747.
- Ioannidis, J. P., Munafo, M. R., Fusar-Poll, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences: Detection, prevalence, prevention. *Trends in Cognitive Sciences*, *18*, 235–241.
- Johnston, C., Lahey, B. B., & Matthys, W. (2013). Editorial policy for candidate gene studies. *Journal of Abnormal Child Psychology*, *41*, 511–514.
- Keller, M. C. (2014). Gene \times environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, *75*, 18–24.
- Levesque, R. J. R. (2015). Statistical guidelines for publishing in the Journal of Youth and Adolescence. *Journal of Youth and Adolescence*, *44*, 2391–2394.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the Education Sciences. *Educational Researcher*, *43*, 304–316.
- McGaghie, W. C. (2015). Varieties of integrative scholarship: Why rules of evidence, criteria, and standards matter. *Academic Medicine*, *90*, 294–302.
- Mueller, K. F., Briel, M., Strech, D., Meerpohl, J. J., Lang, B., Motschall, E., et al. (2014). Dissemination Bias in Systematic Reviews of Animal Research: A Systematic Review. *PLoS ONE*, *9*(12), e116016.
- Mullineaux, P. Y., & DiLalla, L. F. (2015). Genetic influences on peer and family relationships across adolescent development: Introduction to the special issue. *Journal of Youth and Adolescence*, *44*, 1347–1359.
- Nordmann, A. J., Kasenda, B., & Briel, M. (2012). Meta-analyses: What they can and cannot do. *Swiss Medical Weekly*, *142*, w13518.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, 943–951.
- Paré, G., Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, *52*, 183–199.
- Pashler, H., & Harris, C. R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, *7*, 531–536.
- Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis

- of confidence? *Perspectives on Psychological Science*, 7, 528–530.
- Risch, N., Herrell, R., Lehner, T., Liang, K. Y., Eaves, L., Hoh, J., et al. (2009). Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: A meta-analysis. *Journal of the American Medical Association*, 301, 2462–2471.
- Salvatore, J. E., & Dick, D. M. (2015). Gene–environment interplay: Where we are, where we are going. *Journal of Marriage and Family*, 77, 344–350.
- Schlomer, G. L., Cleveland, H. H., Vandenberg, D. J., Fosco, G. M., & Feinberg, M. E. (2015). Looking forward in candidate gene research: Concerns and suggestions. *Journal of Marriage and Family*, 77, 351–354.
- Shuster, L. I., & Cottrill, C. (2015). Ease of articulation: A replication. *Journal of Communication Disorders*, 56, 1–7.
- Simons, D. J. (2014). The value of direct replication. *Perspectives on Psychological Science*, 9, 76–80.
- Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, 9, 59–71.
- Yzerbyt, V. Y., Muller, D., & Judd, C. M. (2004). Adjusting researchers' approach to adjustment: On the use of covariates when testing interactions. *Journal of Experimental Social Psychology*, 40, 424–431.