

Mixed-Methods Systematic Reviews: Integrating Quantitative and Qualitative Findings

Angela Harden, PhD, Professor of Community and Family Health
University of East London

Dr. Angela Harden is a social scientist with expertise in public health and evidence-informed policy and practice. She has conducted extensive research into the health of young people and the communities in which they live. Dr. Harden has a keen interest in research synthesis and knowledge translation. She is widely known for her methodological work integrating qualitative research into systematic reviews and is also an active contributor to the Cochrane and Campbell Collaborations. This issue of FOCUS is adapted from Dr. Harden's speech at the "National Institute on Disability and Rehabilitation Research (NIDRR) Knowledge Translation Conference," July 29, 2009, in Washington, DC.

The topic of mixed-methods systematic reviews arises directly from engaging with decision makers to try to produce more relevant research. Although systematic reviews are a key method for closing the gap between research and practice, they have not always proved to be that useful. Too often, the conclusion of systematic reviews is that there is not enough evidence, or not enough good-quality evidence, to answer the research question or inform policy and practice. The work being done with mixed-methods reviews is an effort to address this issue and make systematic reviews more relevant. By including other forms of evidence from different types of research, mixed-methods reviews try to maximize the findings—and the ability of those findings to inform policy and practice.

My basic argument is that integrating qualitative evidence into a systematic review can enhance its utility and impact. To present this argument, I first discuss the key features of a systematic review and

the use of reviews with different types of studies. I then introduce a framework for conducting mixed-methods systematic reviews and provide an example to make the conceptual ideas more concrete. The framework I present has been developed over several years by myself and colleagues at the Evidence for Policy and Practice Information and Co-ordinating Centre (the EPPI-Centre), Social Science Research Unit, Institute of Education, University of London. My presentation draws on several publications we have written on these types of reviews (Harden et al., 2004; Harden & Thomas, 2005; Oliver et al., 2005; Thomas et al., 2004; Thomas & Harden, 2008).

Key Features of a Systematic Review

In general, the goal of a systematic review is to bring together all existing research studies focused on a specific question or intervention as a shortcut to the literature. Specifically, a systematic review integrates and interprets the studies' findings; it is not just a list

of studies. In other words, a well-done systematic review does something with the findings to increase understanding. Furthermore, a systematic review is a piece of research—it follows standard methods and stages. In marketing systematic reviews it is often useful to emphasize how and why they are pieces of research in their own right rather than just literature reviews.

Another common misconception is that systematic reviews do not require extra time and money because they are something you are supposed to be doing already—you are supposed to be reviewing the literature, and a systematic review is just a way to do so more quickly. I am constantly having to dispute this misconception. Conducting a systematic review requires a significant investment in money, time, resources, and personnel. Like other research, a systematic review requires following specific steps to minimize bias, the introduction of errors, and the possibility of drawing the wrong conclusion. This process includes trying to minimize the bias in individual studies, so a crucial stage of a systematic review involves assessing the quality of each study. These steps are part of the methodology used in the review to search for studies and assess their quality. And like other types of research, systematic reviews should describe their methodology in detail for transparency.

Traditional Systematic Reviews. Most of you will be familiar with the stages of a “traditional” systematic review: starting with developing the review questions and protocol; defining the inclusion criteria for studies; searching exhaustively for published and unpublished studies; selecting studies and assessing their quality; synthesizing the data; and then analyzing, presenting, and interpreting the results. This typical approach is linear in format; but, in reality, conducting a systematic review is a more iterative and circular process. Moreover, although I use a traditional systematic review here as an example, I do not think there really is such a thing.

Figure 1 illustrates what you may see as the product of a typical systematic review of trials (DiCenso,

Guyatt, Willan, & Griffith, 2002). The figure depicts the forest plot from a published review of research on the effectiveness of interventions to reduce teenage pregnancy in the United States and Canada. In this typical forest plot, the small dots represent the effect sizes of the studies, which indicate whether the interventions made a difference. The horizontal lines passing through the dots represent the confidence intervals around the effect sizes. The vertical line running down the middle of the forest plot represents the line of no effect. Thus, the dots that appear in the area to the left of the vertical line, which favors the intervention, suggest the interventions did reduce teenage pregnancy. The dots that appear in the area to the right of the line, which favors the control, suggest the interventions had no effect.

The ultimate goal of this type of systematic review is to produce an overall pooled estimate, which appears at the bottom of the graph in Figure 1. This pooled estimate indicates that the interventions had no effect on teenage pregnancy. The types of interventions reviewed in the studies were school-based sex education, family planning clinics, and abstinence programs—all programs with somewhat of a bias toward sex education. Although these types of interventions did have a positive effect on increasing knowledge and promoting more positive attitudes toward using contraceptives, in the end they did not reduce teenage pregnancy. This lack of effect may be because the interventions did not target all the determinants of teenage pregnancy. For example, there is increasing evidence that social disadvantage is highly correlated with teenage pregnancy, and the interventions in these studies did not really address social disadvantage.

Policy and Practice Questions in Systematic Reviews. Figure 1 represents one type of systematic review—one that uses statistical meta-analysis to synthesize the effect sizes of randomized controlled trials and then provides a forest plot to show the overall pooled effect. In fact, this type of review is so typical that it has practically become synonymous with systematic reviews.

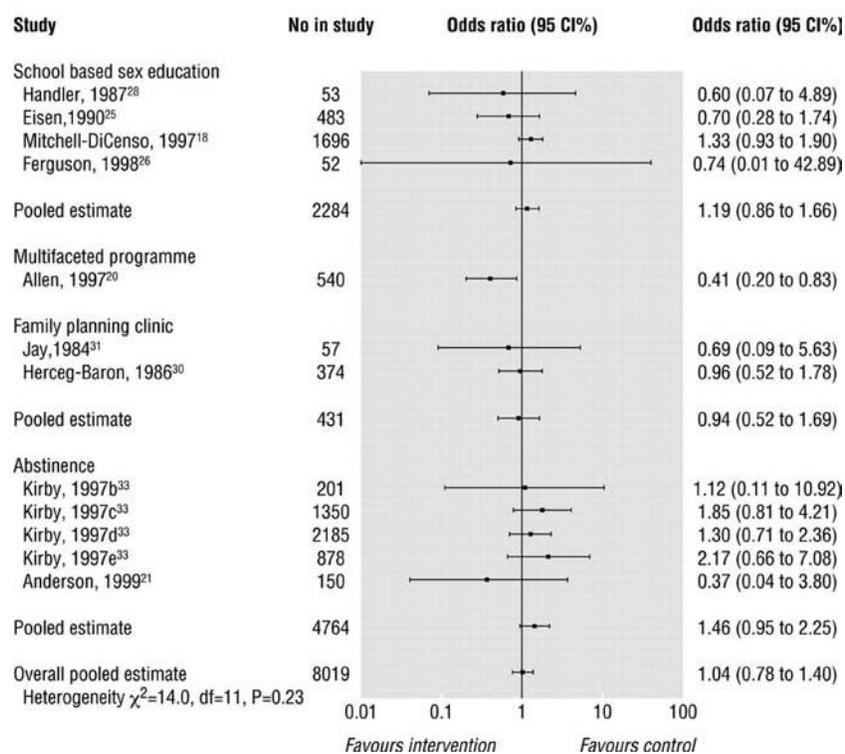


Figure 1: Effectiveness of Interventions to Reduce Teenage Pregnancy in Canada and the United States

Note. Reproduced from “Interventions to Reduce Unintended Pregnancies Among Adolescents: Systematic Review of Randomised Controlled Trials,” by A. DiCenso, G. Guyatt, A. Willan, and L. Griffith, 2002, *BMJ*, 324(7531), pp. 1426–1430. Copyright 2002 BMJ Publishing Group Ltd. Reprinted with permission from the publisher.

This typical systematic review focuses on the question of effectiveness, meaning “what interventions work.” However, the example highlights how systematic reviews—especially in areas, like public health, that involve social interventions rather than medical, surgical, or drug interventions—may need to look at policy and practice concerns before or alongside questions of effectiveness. For instance, we may want to look at the nature of a problem before we start to plan interventions. Later, when evaluating the interventions, we may want to know not only *what* works but also *why* it works and *for whom*.

The question of effectiveness is only one of a range of policy and practice questions. Others include questions about screening and diagnosis, risk and protective factors, meanings and process, economic issues, and issues of methodology. In the emerging field of systematic reviews, the emphasis so far has focused mainly on developing methods around effectiveness. Much less work is being done around other types of questions. Thus, the field remains

wide-open, offering researchers the potential to do a variety of interesting and exciting work.

Systematic Reviews and Diverse Study Types

Because we have a range of policy and practice questions, we need to use different types of research findings to answer them. Likewise, different types of research findings are going to require different types of synthesis. For this reason, a statistical meta-analysis may not fit all types of research.

Statistical Meta-Analysis. As the previous example illustrates, the underlying logic with statistical meta-analysis is aggregation. It is a pooling mechanism. We take the effect sizes from trials, and we pool them to try to see what the overall effect is. We transform the findings from individual studies onto a common scale. We explore variation, we pool the findings, and then display the synthesis in a graphic form.

Meta-Ethnography. Another type of synthesis is meta-ethnography. Educational researchers

developed this approach in the early 1980s (Noblit & Hare, 1988). More recently, health services researchers have started to explore its applicability for synthesizing qualitative research within health care. Unlike meta-analysis, the underlying logic of meta-ethnography is interpretation rather than aggregation. In other words, although meta-ethnography includes an element of aggregation, its synthesis function is interpretation. Rather than pooling findings from studies, key concepts are translated within and across studies, and the synthesis product is a new interpretation. In a meta-ethnography, the synthesis goal is to achieve a greater level of understanding or conceptual development than can be found in any individual empirical study.

For example, a meta-ethnography has been conducted to synthesize research on teenage pregnancy and the experiences of teenage mothers in the United Kingdom (McDermott & Graham, 2005). The synthesis provided a different type of product, consisting of recurring concepts and themes across studies, like poverty, stigma, resilient mothering, the good mother identity, kin relations and social support, and the prioritization of the mother-child relationship.

The meta-ethnography involved two main stages. The first stage was to draw out the concepts from the individual studies. The next stage was to formulate a new interpretation that integrated those concepts across studies into a line of argument. The interpretation of this synthesis was that teenage mothers in the United Kingdom have to mother under very difficult circumstances. These young women not only often have to mother in poverty but also are positioned outside of the boundaries of normal motherhood. Therefore, teenage mothers in the United Kingdom have to draw on the only two resources available to them: their own personal resources and their kin relationships. What you find in teenage mothers' accounts is a real prioritization of the mother-child relationship—a strong investment in the good mother identity.

This finding has implications for policy and practice in the United Kingdom, where the strategy has focused on social exclusion as the cause and consequence of teenage pregnancy—a strategy that presents teenage mothers as feckless individuals. In contrast, the finding of the meta-ethnography indicates that teenage motherhood can actually be a route into social inclusion rather than social exclusion because of the emphasis on kin relationships, family support, and the good mother identity.

Combining Qualitative and Quantitative Findings in the Same Review

The two teenage pregnancy examples each use a different type of synthesis—the first uses a meta-analysis of interventions; the second uses meta-ethnography. The question is how can we combine these two methods in one systematic review.

In the mixed-methods systematic reviews I have been conducting with colleagues at the EPPI-Centre, there are three ways in which the reviews are mixed:

1. The types of studies included in the review are mixed; hence, the types of findings to be synthesized are mixed.
2. The synthesis methods used in the review are mixed—statistical meta-analysis and qualitative.
3. The review uses two modes of analysis—theory building and theory testing.

As an example, I am going to focus on one systematic review about children and healthy eating, commissioned by the English Department of Health (Thomas et al., 2003). Around the time the review was commissioned, health policy in the United Kingdom included three main priorities in terms of public health: the promotion of healthy eating, physical activity, and mental health.

We worked closely with colleagues at the Department of Health to come up with the review questions. The central questions were the following:

1. What is known about the barriers to and facilitators of healthy eating among children?
2. Do interventions promote healthy eating among children?

3. What are children’s own perspectives on healthy eating? (There was a great interest in the department about not only interventions but also children’s own views and what we can learn from them.)
4. What are the implications for intervention development?

In our review process (see Figure 2), we searched, screened, and mapped the research. And we were overwhelmed by the number of studies we found. With systematic reviews, you often think you have a narrow enough question, that there cannot possibly be that much out there to find. However, when you actually start to search systematically, you uncover thousands and thousands of references.

To help narrow our search, we built into our process a two-stage procedure for consulting with policymakers to see which area of the literature they wanted to focus on. During that time in the United Kingdom, there was a big push to get people to eat more fruits and vegetables, such as the Five-a-Day campaign. In keeping with this push, we narrowed our focus to fruit and vegetable consumption. Within this focus, we had three syntheses. First, we looked at trials of interventions to promote fruit and vegetable consumption. Second, we looked at qualitative studies on children’s own perspectives

and experiences on eating fruit and vegetables and on healthy eating in general. Third, we looked at an integration of the two types of syntheses. The result was a single systematic review with three syntheses.

As Figure 2 illustrates, we conducted the first two syntheses and then used them to create Synthesis 3. Throughout the process, we applied the same principles across the studies but used different methods for each type. The first stage was quality assessment. For the trials, we looked at pre- and post-data on outcomes to be reported. We wanted to see all the data on the outcomes and on an equivalent control or comparison group. For the qualitative studies of children’s views, we looked at the quality of the reporting and at the strategies used to enhance rigor, data analysis, and collection. The next stage was data extraction. For this stage, we used a standard protocol that varied by type of study to capture different types of data—numerical, categorical, or textual. That is, we did not ask the same questions of the trials as we did of the qualitative studies. In the final stage, to synthesize the data, we conducted a statistical meta-analysis to pull the effect sizes from the trials in Synthesis 1. Next, we used a thematic analysis to synthesize the findings of the qualitative studies in Synthesis 2. We then integrated the two types of findings by using

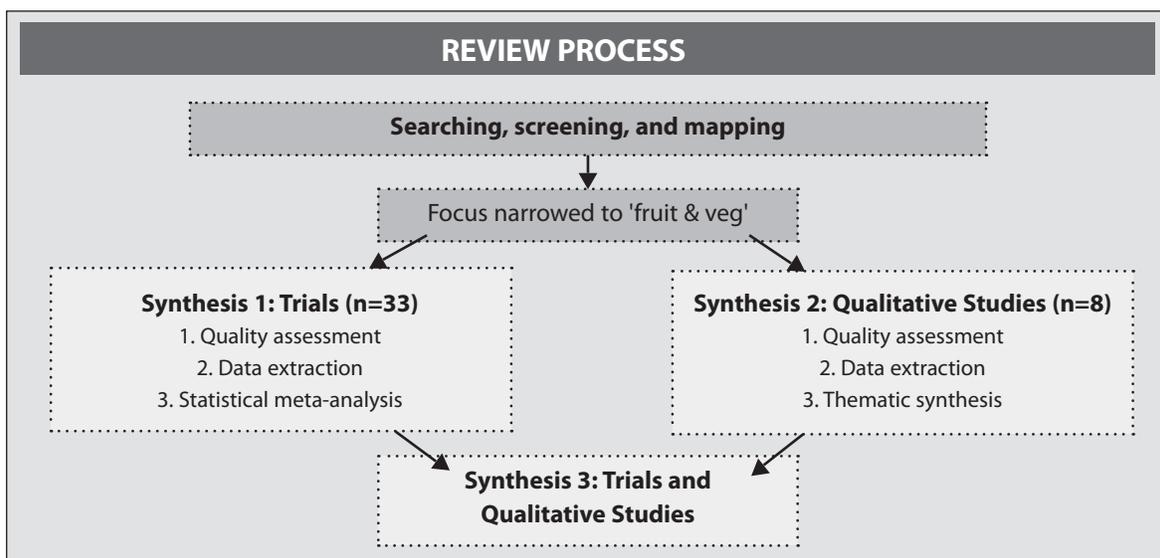


Figure 2: Sample Process for a Mixed-Methods Systematic Review

Synthesis 2 to interrogate Synthesis 1, producing Synthesis 3.

Because part of the review was about trying to privilege children's own perspectives and learn from them, we also assessed the extent to which study findings were rooted in those perspectives. In terms of the trials, we used conventional statistical meta-analysis on six different outcome measures. We looked at increases in fruit consumption alone, at increases in knowledge about fruit and vegetables, and at heterogeneity across studies. To analyze these studies, we used sub-group analysis and also did a qualitative analysis of the textual data from the trials.

For the first synthesis, we produced a forest plot showing the results of each of the individual trials we included. The graph indicates that the interventions did have a small positive effect on increasing children's fruit and vegetable intake. The overall effect size from the trials was an increase of about a quarter of a portion of fruit and vegetables a day. The individual studies varied widely in their findings, however, with some showing large increases and others actually showing decreases in fruit and vegetable intake. This variation raised questions about effectiveness and why one intervention may be more effective than another.

In the second synthesis, we wanted to integrate the children's perspectives on healthy eating. Were the trials really targeting what children were saying was important in terms of encouraging them to eat fruits and vegetables? For this synthesis, we used the steps of thematic analysis, meaning we had a stage where we were coding text and developing descriptive themes. We first conducted a line-by-line coding of the text. Next, we grouped the initial codes and collapsed codes. We then had a third stage where we generated more explanatory and analytical themes. Figure 3 provides our list of 13 descriptive themes, grouped according to children's understandings of healthy eating and influences on food eaten.

The descriptive themes capture several ideas: children label good and bad foods, they are aware of some of the health consequences of food, they

have clear food preferences, and they like to eat to socialize. Children also are aware of some of the food contradictions in school, such as how teachers talk about eating healthfully, while the foods provided in the school canteen are not particularly healthful. In addition, children talked about "food rules"—the rules their parents enforce about food, like not being able to leave the table until they have eaten all their vegetables—and then talked with relish about breaking those rules—when the parents go out, the sweets come out.

In the second stage of Synthesis 2, we wanted to push this set of descriptive themes further to try to explain what was going on around healthy eating and children's consumption of fruits and vegetables. We identified a set of six analytical themes. For example, the first theme is that children do not necessarily see their role as being interested in health. They consider taste, not health, to be a key influence on their food choices. As a result, food labeled as healthy may lead some children to reject it. "I don't like it, so it must be healthy." Children did not like thick, green vegetables that are supposed to be full of vitamins, but they did like sweet corn and carrots—that is, small vegetables that are sweet tasting. Not all children thought this way, however. Some of the girls, especially, were influenced by health when making food choices. Children also did not see buying healthy foods as a legitimate use of their pocket money; they thought their parents should be buying the oranges and apples. Children instead wanted to buy sweets with their pocket money.

So how did we integrate the two syntheses? From the analytical themes, we came up with a set of recommendations for interventions that reflected children's views. We then used our recommendations to interrogate the interventions evaluated in the trials. We were interested in how well those interventions matched our recommendations.

For the first theme—children do not see their role as being interested in health and do not see future health consequences as personally relevant or credible—we recommended reducing the health emphasis of messages and branding fruits and

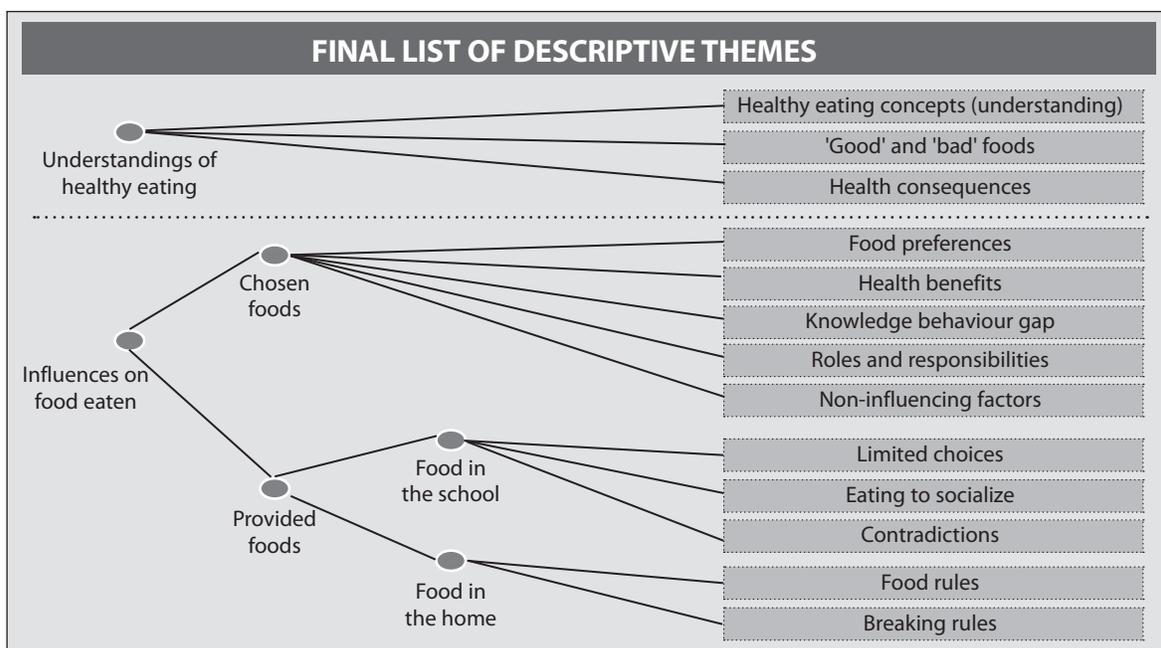


Figure 3: Descriptive Themes: U.K. Children's Perspectives on Healthy Eating

vegetables as tasty rather than healthy. The analytical themes also indicated that fruit, vegetables, and sweets have very different meanings for children. The Five-a-Day campaign does not really recognize that view. The campaign calls for five pieces of fruit or vegetables a day and does not distinguish between the two. On the basis of the first theme, a natural recommendation or implication for interventions is not to promote fruit and vegetables in the same way within the same intervention. We did not find any trials that tested this approach. We did, however, have a good pool of trials that had looked at branding fruits and vegetables as exciting and tasty as well as several trials that focused on reducing the health emphasis in messages. Looking across studies and considering the effect sizes, the trials with the biggest effect on children's increase in vegetable intake had little or no emphasis on health messages. This finding reiterates that children's views really do matter; their views have implications for policy and practice.

This question of combining qualitative and quantitative studies does pose a risk of recreating the paradigm wars within systematic reviews, something I am keen to prevent. At the same time, I see the use of mixed-methods reviews as a growing trend in the literature. For many people conducting meta-

ethnographies, their point of departure is statistical meta-analysis. This group is trying to show how we need a different model to review qualitative research.

Mixed-methods systematic reviews can be defined as combining the findings of "qualitative" and "quantitative" studies within a single systematic review to address the same overlapping or complementary review questions. You notice I have scare quotes around the terms *qualitative* and *quantitative*. In this definition, the two terms are useful heuristic devices to signify broadly what we mean by different types of research, but I am not completely convinced of the value of these labels. That is, I am not wholly convinced it is a good idea to describe studies as either qualitative or quantitative because I think every study has aspects of both.

Summary

Including diverse forms of evidence is one way to increase the relevance of systematic reviews for decision makers. In the previous case, we had a number of trials that had looked at the question of "how effective are interventions to promote fruit and vegetable intake." The trials showed a huge heterogeneity, and the inclusion of qualitative research helped us to explain some of that

heterogeneity. Including qualitative research also helped us to identify research gaps. We did not find any trials of interventions that promote fruit and vegetables in the same way, like the Five-a-Day campaign.

Using a mixed-methods model is one way to answer a number of questions in the same systematic review. Rarely do decision makers have just one question to answer; they are more likely to have a series of questions. The mixed-methods model enables us to integrate quantitative estimates of benefit and harm with more qualitative understanding from people's lives. This integration helps determine not only the effects of interventions but also their appropriateness. This concept is similar to that of social validity.

What is really important to me about mixed-methods design is that it facilitates this critical analysis of

interventions from the point of view of the people the interventions are targeting. This design brings their experience to bear and draws on their different skills and expertise. Another feature of the mixed-methods design is that it preserves the integrity of the findings of the different types of studies. We are not converting qualitative findings into numbers or quantitative findings into words. The technique uses complementary frameworks for qualitative and quantitative research to preserve each method. The fruit and vegetable systematic review is only one example of a whole series of mixed-methods reviews using the same approach. To see some other examples, go to the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) Web site at <http://eppi.ioe.ac.uk/EPPIWeb/home.aspx>.

Acknowledgment: I would like to acknowledge my former colleagues at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; the English Department of Health for funding the program of review work in which the mixed methods were developed; and the Economic and Social Research Council (ESRC). I can be reached at a.harden@uel.ac.uk.

References

DiCenso, A., Guyatt, G., Willan, A., & Griffith, L. (2002). Interventions to reduce unintended pregnancies among adolescents: Systematic review of randomized controlled trials. *BMJ*, *324*(7531), 1426–1430.

Harden, A., Garcia, J., Oliver, S., Rees, R., Shepherd, J., Brunton, G., et al. (2004). Applying systematic review methods to studies of people's views: An example from public health. *Journal of Epidemiology and Community Health*, *58*, 794–800.

Harden, A., & Thomas, J. (2005). Methodological issues in combining diverse study types in systematic reviews. *International Journal of Social Research Methodology*, *8*(3), 257–271.

McDermott, E., & Graham, H. (2005). Resilient young mothering: Social inequalities, late modernity and the 'problem' of 'teenage' motherhood. *Journal of Youth Studies*, *8*(1), 59–79.

Noblit, G. W., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. Newbury Park, CA: Sage Publications.

Oliver, S., Harden, A., Rees, R., Shepherd, J., Brunton, G., Garcia, J., et al. (2005). An emerging framework for integrating different types of evidence in systematic reviews for public policy. *Evaluation*, *11*(4), 428–446.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*, 45. doi:10.1186/1471-2288-8-45

Thomas, J., Harden, A., Oakley, A., Oliver, S., Sutcliffe, K., Rees, R., et al. (2004). Integrating qualitative research with trials in systematic reviews: An example from public health. *BMJ*, *328*, 1010–1012.

Thomas, J., Sutcliffe, K., Harden, A., Oakley, A., Oliver, S., Rees, R., et al. (2003). *Children and healthy eating: A systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Copyright © 2010 by SEDL

FOCUS: A Technical Brief From the National Center for the Dissemination of Disability Research was produced by the National Center for the Dissemination of Disability Research (NCD DR) under grant H133A060028 from the National Institute on Disability and Rehabilitation Research (NIDRR) in the U.S. Department of Education's Office of Special Education and Rehabilitative Services (OSERS). NCD DR is funded 100% by NIDRR at \$750,000 per project year.

NATIONAL CENTER FOR THE DISSEMINATION OF DISABILITY RESEARCH

Advancing Research, Improving Education



NCD DR's scope of work focuses on developing systems for applying rigorous standards of evidence in describing, assessing, and disseminating outcomes from research and development sponsored by NIDRR. The NCD DR promotes movement of disability research results into evidence-based instruments such as systematic reviews as well as consumer-oriented information systems and applications.

SEDL operates the NCD DR. SEDL is an equal employment opportunity/affirmative action employer and is committed to affording equal employment opportunities for all individuals in all employment matters. Neither SEDL nor the NCD DR discriminates on the basis of age, sex, race, color, creed, religion, national origin, sexual orientation, marital or veteran status, or the presence of a disability. The contents of this document do not necessarily represent the policy of the U.S. Department of Education, and you should not assume endorsement by the federal government.