

Health Behavior Interventions Literature Review

Jamie Harris

Researcher

July 24, 2020

© 2020 Sentience Institute

Edited by Jacy Reese Anthbis and Kelly Anthbis. Many thanks to Juliette Finetti for providing advice and feedback.

Abstract

There is an extensive research literature on the effectiveness of various interventions for changing health behaviors, such as dietary habits, physical activity, and smoking. Existing reviews and overviews tend to aggregate research on the effects of specific interventions on particular health behaviors. This review provides a largely qualitative, non-statistical summary of the evidence for the effectiveness of intervention types across health behaviors, by aggregating 706 research items (mostly systematic reviews and meta-analyses) based on the strength of evidence. This review makes a number of subjective judgment calls and uses novel methods of evaluation in order to quickly digest and summarize an extremely large evidence base. It focuses on applications to the farmed animal movement, especially to the behavioral change of reducing animal product consumption. In general, while the health behavior literature includes a very large number of studies, there is much inconsistency in wording, methodology, and subject matter, which makes it difficult to extract useful insights for behavior change advocates. However, some conclusions are warranted. Key findings include that almost all types of health behavior interventions targeted at individuals or small groups seem likely to have effect sizes conventionally interpreted as “small” or “very small,” that their effect sizes tend to be even smaller in the long term, and that interventions with educational and behavioral components outperform solely educational interventions.



Table of Contents

Abstract	1
Table of Contents	2
Introduction	2
Summary of strategic implications	4
Appendices and supplementary spreadsheets	6
Methodology	7
Search strategy	8
Explanation and discussion of the four qualitative analyses used in this literature review	11
“Overall judgement” and “Strength of evidence assessments”	12
“Size of effects or differences shown through direct tests”	14
“Size of differences shown through indirect comparisons”	17
Moderator analyses	18
Results and discussion	19
Strengths and weaknesses of the health behavior literature	32
Strengths of the health behavior literature	32
Weaknesses of the health behavior literature	33
Limitations	35
Potential Items for Further Study	37
Bibliography	38

Introduction

Health behaviors are “overt behavioral patterns, actions and habits that relate to health maintenance, to health restoration and to health improvement.”¹ They are often associated with lifestyle features that affect health and lifespan, such as diet and nutrition, physical activity (henceforth PA), smoking, alcohol consumption, sexual health and contraceptive use, adherence to medication, and vaccination, immunization, screening, and appointment attendance.² Occasionally the terms “health behavior” and “risk behavior” are used separately to refer to actions that improve or endanger health respectively, although in this review, the term “health behavior” is used to refer to both health and risk behaviors. The study of health behavior is a vast subject area, drawing not only on medical science but also on social sciences such as behavioral economics and psychology.

While we are generally interested in what the health behavior literature can tell us about changing behavior, our particular focus is on animal product consumption and other behaviors of interest to the farmed animal movement, such as participating in demonstrations against factory farming, though animal product consumption seems most similar to typical health behavior changes. We decided to write this literature review because, as an international and well-funded field, health behavior seemed likely to have a fuller evidence base, and perhaps more certain conclusions, than exist currently in the farmed animal movement. Of course, there are many differences between the farmed animal and health behavior contexts which mean that conclusions in one field will not automatically transfer across to the other. However, if some features, trends, or effects are consistent across several different health behaviors, then it seems more likely that these findings will transfer to the farmed animal context.

¹ David S. Gochman (ed.), *Handbook of health behavior research, vol. 1* (New York: Plenum, 1997), 3. On 7, Gochman adds that “[h]ealth behavior is conceptually distinct from treatment and from physiological/biological/pharmacological responses to treatment. It is also conceptually distinct from health care and from the organization or structure of the health care delivery system.”

² Nedra B. Belloc and Lester Breslow, “Relationship of physical health status and health practices,” *Preventive Medicine* 1, no. 3 (1972), 409–421, for example, found that “common health practices, including hours of sleep, regularity of meals, physical activity, smoking and drinking” were associated with “physical health status,” as measured by “[a] physical health spectrum” that was developed from answers to questions about disability, chronic conditions, impairments, symptoms, and energy level.” This was “independent of age, sex, and economic status.”

David S. Gochman (ed.), *Handbook of health behavior research, vol. 1* (New York: Plenum, 1997), 4, after listing several behaviors which can prevent ill-health, notes that, “[a] survey of physicians (Sobal, Valente, Muncie, Levine, & DeForge, 1985) found strong professional consensus about the preventive importance of these everyday behaviors, particularly about eliminating smoking and using protective equipment and clothing.” On page 5, Gochman adds that, “health protective behaviors” are also important.

Summarizing more recent evidence, Susan Michie and Caroline E. Wood, “Health behaviour change techniques,” Mark Conner and Paul Norman (eds.), *Predicting and changing health behaviour: Research and practice with social cognition models* (Maidenhead: Open University Press, 2015; first published 1995), 358 note that, “[p]reventable behaviours, such as smoking, physical activity, unhealthy eating habits, and excessive alcohol consumption have been identified as leading causes of morbidity and mortality in resource-rich countries.”

The division of categories in this review was chosen based on the divisions found most commonly in the literature.

External findings from the health behavior field can therefore provide weak evidence for animal advocacy strategy. Although this review draws strategic implications from the health behavior literature, these implications are only valid insofar as we place weight on external evidence, and should always be interpreted in the context of other available evidence.³

Summary of strategic implications

Below are a number of strategic claims supported by the evidence in this review:

- In the health behavior literature, almost all types of health behavior interventions seem likely to have effect sizes conventionally interpreted as “small” or “very small.” The overall differences between intervention types seem minor, despite substantial variance between individual studies or meta-analyses, so animal advocates should usually not expect large differences between intervention types without additional evidence.
- Incentives, price changes, bans on undesired behavior, and other forms of policy change seem to have larger effect sizes than most individual or small group interventions.
- A few techniques have correlational evidence for their effectiveness. These include the provision of instruction, prompting of specific goal setting, prompting of self-monitoring of behavior, prompting of practice, facilitation of social comparison, barrier identification or problem solving, social support, and interpersonal skills training.⁴
- Indirect comparisons between studies suggest that interventions with educational and behavioral components outperform solely educational interventions. Specific behavioral components that show promise include action planning, implementation intentions, goal setting, and motivational interviewing.⁵
- Although various forms of mass media interventions are frequently used in the farmed animal movement, the evidence for the effectiveness of such tactics in the health behavior literature is surprisingly weak.
- There is little evidence that bans on the advertising of unhealthy products are effective in reducing consumption.
- Reviews tend to find evidence that “nudges” can be effective at changing behavior, though effect sizes are probably small. There is stronger evidence that modifying the size of portions, packaging, or tableware affects consumption than there is for other types of nudges.
- There is relatively strong evidence that interventions delivered remotely, such as through telephones, text messages, or computers, can be effective at changing behavior, though as with other interventions, overall effect sizes tend to be “small” or “very small.” Differences in behavioral effects between face-to-face interventions and remote delivery are often not statistically significant, though

³ “[meta] Social movements vs. EAA randomized controlled trials (RCTs) vs. intuition/speculation/anecdotes vs. external findings” in “Summary of Evidence for Foundational Questions in Effective Animal Advocacy,” Sentience Institute, last updated June, 2018, [https://www.sentienceinstitute.org/foundational-questions-summaries#\[meta\]-social-movements-vs.-eaa-randomized-controlled-trials-\(rcts\)-vs.-intuition/speculation/anecdotes-vs.-external-findings](https://www.sentienceinstitute.org/foundational-questions-summaries#[meta]-social-movements-vs.-eaa-randomized-controlled-trials-(rcts)-vs.-intuition/speculation/anecdotes-vs.-external-findings).

⁴ Some of these intervention types also have experimental evidence for their effectiveness.

⁵ For definitions of these terms, see [Appendix A: Definitions and Discussion by Intervention Type](#).

indirect comparison between the different effect size estimates within meta-analyses suggests that face-to-face interventions have slightly larger behavioral effects than otherwise similar non face-to-face interventions.

- Contrary to expectations, interventions delivered in group format seem to have very slightly larger behavioral effects than individual interventions.
- There is little evidence that the effect of behavior change interventions depends on particular settings such as schools, family homes, or workplaces.
- There is some evidence that “brief interventions” can be effective and have small effect sizes (as other health behavior interventions do). Additionally, there is some evidence that brief interventions are not made more effective by increasing the amount of contact time involved in interactions.
- There is some evidence that the distribution of print-based materials can positively affect behavior, but one review found negative effects on motivation.
- Broadly, the health behavior literature provides very mixed evidence regarding whether increases in the intensity of interventions — such as through greater amounts of contact time, increases in the number of contacts, or lengthened duration of the intervention — are likely to increase effectiveness. There is some evidence that increasing the number of contacts in interventions delivering tailored print-based materials increases their effectiveness, but there is surprisingly consistent evidence that increasing the intensity of online and computer-based interventions does not increase their effectiveness and some evidence that this actually reduces their effectiveness.
- The evidence relating to increases in complexity (such as incorporating a larger number of behavior change techniques or carrying out an intervention in multiple settings) is mostly positive, with meta-analyses suggesting moderate to large increases in intervention effectiveness at causing behavioral change.
- In aggregate, meta-analyses suggest moderate increases in intervention effectiveness from the use of tailoring (that is, where content is matched to an individual’s characteristics or needs).
- There was fairly consistent evidence that taxes, subsidies, and other interventions affecting price had greater effects on people with low socioeconomic status (SES). There was evidence that education or information only interventions are less effective among people with low SES.
- The evidence relating to the long-term effects of interventions is far weaker than expected.
- One-by-one diet interventions seem very unlikely to have moderate-or-large effects after one year; from the reviewed interventions, only “direct financial incentives” were found to have anything other than small, very small, or no effects at long-term follow-up. For all interventions where evidence was available, except school-based interventions, meta-analyses suggest that effect sizes are smaller when measured at longer follow-up. In aggregate, indirect comparisons between the different effect size estimates within 17 meta-analyses suggest that behavioral effects are slightly or moderately smaller when measured at follow-up points of a year or more post-intervention than when measured at shorter follow-up points.
- Weaknesses of the health behavior literature, despite decades of research and huge amounts of funding, suggest serious limitations of experimental and observational research in other contexts, such as the farmed animal movement. This makes other types of evidence, such as social movement case studies, relatively more promising.
- There is evidence that changing audience’s intentions changes their behavior, but the behavioral effect sizes tend to be smaller; meta-analyses have found effects as much as 62% smaller, though a

discount rate between 25% and 50% seems more representative for converting intentional effect size to behavioral effect size.

- Although this review focused on behavioral outcome measures, rather than indirect measures such as intentions or knowledge, some intervention types, such as those focused on education, either directly to individuals or via mass media, performed much better on indirect outcome measurements. This suggests that these interventions could have important effects; their impact on behavioral outcomes might only be detected with longer follow-up measurement, or if they are combined with other interventions. Such interventions could be necessary for long-term behavioral change but not sufficient to cause change without the support of other interventions.
- Several social cognition models used in the health behavior literature — especially the Transtheoretical Model⁶ — seem applicable to the farmed animal movement and may help to support intervention design.

Appendices and supplementary spreadsheets

Two appendices provide additional detail and discussion:

- [A: Definitions and Discussion by Intervention Type](#)

⁶ See, for example, Elisabeth Mendes, “An Application of the Transtheoretical Model to Becoming Vegan,” *Social Work in Public Health* 28, no. 2 (2013), 142-9.

L. Wolfenden, N. Nathan, C. M. Williams, “Computer-tailored interventions to facilitate health behavioural Change,” *British Journal of Sports Medicine* 49, no. 22 (November 2015), 1478-9 found very small significant effects overall and note that, “[t]ailored interventions based on assessments of participants prior to each intervention contact (dynamic tailoring) were more effective than interventions tailored based on a single assessment (static tailoring) at baseline (g=0.19 vs 0.14, p=0.01).”

However, three reviews provide evidence that suggests that even if tailoring is effective, repeated variations in tailored content by the stage of change has no effect on behavioral outcomes.

Mia Liza A. Lustria, Seth M. Noar, Juliann Cortese, Stephanie K. Van Stee, Robert L. Glueckauf, and Junga Lee, “A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions,” *Journal of Health Communication* 18, no. 9 (June 2013), 1039-69 found that moderator analysis “demonstrated no significant difference among studies that administered multiple assessments over the course of the intervention to create tailored messages (n = 10) compared with those that used a single tailoring assessment at baseline (n = 11).”

Kate Cahill, Tim Lancaster, and Natasha Green, “Stage-based interventions for smoking cessation,” *Cochrane Database of Systematic Reviews* 11 (November 2010) note that, “[f]our trials, which directly compared the same intervention in stage-based and standard versions, found no clear advantage for the staging component. Stage-based versus standard self-help materials (two trials) gave a relative risk (RR) of 0.93 (95% CI 0.62 to 1.39). Stage-based versus standard counselling (two trials) gave a relative risk of 1.00 (95% CI 0.82 to 1.22).” Although there is insufficient information to rate the evidence by GRADE criteria, it seems likely to be either of moderate or low quality.

Jamie Hartmann-Boyce, Tim Lancaster, and Lindsay F. Stead, “Print-based self-help interventions for smoking cessation,” *Cochrane Database of Systematic Reviews* 6 (June 2014; first published 2002) found in studies comparing tailored print-based self-help materials to non-tailored print-based self-help materials that “results favoured tailored interventions when the tailored interventions involved more mailings than the non-tailored interventions” (RR 1.42, 95% CI 1.20 to 1.68, 9 studies), “but not when the two conditions were contact-matched” (RR 1.07, 95% CI 0.89 to 1.30, 10 studies).” They rated the quality of the evidence as moderate for each finding.

- [B: Theory Used in Health Behavior Intervention Design and Research](#)

Narrative summaries of findings by health behavior topic are available upon request, as is a document listing the factors that were considered by the author for assessing the relevance of each of the health behavior areas to the farmed animal movement. Each of these appendices and supplementary documents was less thoroughly edited than Sentience Institute's usual standard, in order to reduce the time required to summarize the extensive health behavior literature.

Additional supplementary spreadsheets provide detail on research items that were reviewed and evaluated to inform the results in Table 1 below:

- [Strength of Evidence Assessments](#)
- [Effect Size Estimates](#)
- [Moderator Analyses](#)
- [Lists of Included and Excluded Research](#)

Methodology

Prior to the literature review, we expected that health behavior research would provide evidence to help us better understand:

- Foundational questions in effective animal advocacy that relate to consumer behavior, such as whether to focus on promoting reducetarianism or veganism,
- The overall effectiveness of certain types of interventions for encouraging behavior change (e.g. servings of meat reduced per dollar spent on advocacy), and
- The gaps in effective animal advocacy research and what should be prioritized for further study.

In order to focus the review on those topics of most interest:

- Where possible, findings about behavior change outcomes are reported, rather than findings about clinical and health outcomes (e.g. blood pressure), what constitutes "good health," and what general theories of behavior change make sense (except insofar as that facilitates our understanding of cost-effective intervention design, which ended up not being very much).
- Findings about how interventions affect less directly important outcomes, such as measures of knowledge or the various determinants of behavior change included in social cognition models, are rarely included.
- At times, the health behavior literature overlaps with other academic topics that may also be of some interest to the farmed animal movement, such as advertising. In order to keep focused on health behavior, this review avoids delving into these topics, but future research could review each such area.

Search strategy

Because the health behavior literature is so vast, this literature review relies primarily on systematic reviews. The focus is on Cochrane reviews where possible, since they have especially strong methodology.⁷ However, Cochrane reviews do not comprehensively cover the health behavior literature.⁸ The search was therefore widened beyond research by the Cochrane Collaboration to consider systematic reviews in other journals. “High-level reviews,” “meta-reviews,” and systematic “reviews of reviews” (henceforth collectively referred to as overviews) are also included. Individual studies are very rarely included.

There are established best-practice methodologies for systematic reviews,⁹ but these are time-consuming procedures and can cost “upwards of \$100,000 each.”¹⁰ A review of “rapid reviews” found that even this more stripped back methodology requires on average 3.2 months to complete for just a small subset of the evidence that this literature review seeks to cover.¹¹ Given the vast breadth of research being evaluated in this literature review and its uncertain transferability to the farmed animal movement, a more efficient method was needed for the health behavior topics considered here. Nevertheless, as discussed below, steps have been taken to avoid arbitrary or biased selection of evidence that might limit the robustness of the results.

⁷ “About Cochrane Reviews,” The Cochrane Collaboration, accessed January 17, 2019 <https://www.cochranelibrary.com/about/about-cochrane-reviews> notes that, “[a] Cochrane Review is a systematic review of research in health care and health policy that is published in the Cochrane Database of Systematic Reviews.” These reviews “use explicit, systematic methods that are selected with a view aimed at minimizing bias, to produce more reliable findings.”

Katherine R. Jones, Nirmala Lekhak, and Napatsawan Kaewluang, “Using Mobile Phones and Short Message Service to Deliver Self-Management Interventions for Chronic Conditions: A Meta-Review,” *Worldviews on Evidence-Based Nursing* 11, no. 2 (April 2014), 83 notes that, “Cochrane systematic reviews are considered the ‘gold standard’ of systematic reviews, with rigorous requirements for their performance and reporting;” it found all 4 Cochrane reviews included to have the maximum AMSTAR (A MeaSurement Tool to Assess systematic Reviews) rating (see “What is AMSTAR,” AMSTAR, accessed January 17, 2019 https://amstar.ca/About_Amstar.php). This literature review therefore initially focused discussion on these results.

Holden Karnofsky, “Surveying the research on a topic,” last updated July 2016, <https://blog.givewell.org/2012/09/06/surveying-the-research-on-a-topic/> of the Charity evaluator GiveWell also considers the Cochrane Collaboration to provide some of the strongest literature reviews.

⁸ For example, Antonio José Grande, Tammy Hoffmann, and Paul Glasziou, “Searching for randomized controlled trials and systematic reviews on exercise. A descriptive study,” *Sao Paulo Medical Journal* 133, no. 2 (April 2015), found that, “[u]p until 2011, 9,354 RCTs about exercise were published in 1,250 journals and 1,262 SRs in 513 journals... The Cochrane Database of Systematic Reviews was the principal source for SRs, with 9.8% of the total.”

⁹ See Julian P. T. Higgins and Sally Green (eds.) “Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0” (March 2011), <http://handbook-5-1.cochrane.org/>.

¹⁰ Holden Karnofsky, “Surveying the research on a topic” last updated July 2016, <https://blog.givewell.org/2012/09/06/surveying-the-research-on-a-topic/>.

¹¹ Ahmed M. Abou-Setta, Maya Jeyaraman, Abdelhamid Attia, Hesham G. Al-Inany, Mauricio Ferri, Mohammed T. Ansari, Chantelle M. Garritty, Kenneth Bond, and Susan L. Norris, “Methods for Developing Evidence Reviews in Short Periods of Time: A Scoping Review,” *PLOS ONE* 11, no. 12 (December 8, 2016), 7.

Initially, broad Google searches were conducted using general search terms such as “health behavior” that identified several recent overviews¹² and the information contained in [Appendix B: Theory Used in Health Behavior Intervention Design and Research](#).

Secondly, Cochrane reviews were searched via the [Cochrane Library](#). Two Cochrane review topics, “Tobacco, Drugs & Alcohol” and “Public Health,” seemed directly relevant to health behavior, and so these were searched comprehensively, checking the title, abstract, and summary of findings for all reviews included in these topic categories. For each review, at least one reference to it was included in this present literature review or it was listed in the spreadsheet “[Lists of included and excluded research](#)” with a brief explanation of why it was excluded.

Many other Cochrane review topics and groups seemed likely to contain some reviews relevant to health behavior. The titles of all reviews in these topics were screened, and references were included to those providing relevant information, though in most cases, the reasons for excluding were not stated, as this seemed unnecessarily time-consuming given their relatively clear irrelevance. Additionally, to check that key reviews had not been missed, several terms were searched in the Cochrane Library and checked in the same manner. The topics, groups, and search terms that were checked in this manner included:

- Consumer & communication strategies (topic)
- Dentistry & oral health (topic; no further reviews included, i.e. none other than those already included from other sources)
- Effective practice & health systems (topic)
- Endocrine & metabolic (topic)
- Heart & circulation (topic)
- Fertility regulation (group)
- Sexually Transmitted Infections (group; no further reviews included)
- Sleep (search term)
- Behaviour (search term)
- Routine (search term)

¹² These reviews were:

Denise de Ridder, Floor Kroese, Catherine Evers, Marieke Adriaanse, and Marleen Gillebaart, “Healthy diet: Health impact, prevalence, correlates, and interventions,” *Psychology & Health* 32 (April 2017), 907–941,

Robert West, “Tobacco smoking: Health impact, prevalence, correlates and interventions,” *Psychology & Health* 32 (August 2017), 1018–1036,

Emmanuel Kuntsche, Sandra Kuntsche, Johannes Thrul, and Gerhard Gmel, “Binge drinking: Health impact, prevalence, correlates, and interventions,” *Psychology & Health* 32 (May 2017), 976–1017,

Ryan E. Rhodes, Ian Janssen, Shannon S. D. Bredin, Darren E. R. Warburton, and Adrian Bauman, “Physical activity: Health impact, prevalence, correlates and interventions,” *Psychology & Health* 32 (May 2017), 942–975,

Elísio Costa, Anna Giardini, Magda Savin, Enrica Menditto, Elaine Lehane, Olga Laosa, Sergio Pecorelli, Alessandro Monaco, and Alessandra Marengoni, “Interventional tools to improve medication adherence: review of literature,” *Patient Preference and Adherence* 9 (September 2015), 1303-1314.

- Adherence (search term in titles)
- Compliance (search term in titles; no further reviews included)
- Physical (search term in titles; no further reviews included)
- Exercise (search term in titles; no further reviews included)
- Diet (search term in titles)
- Nutrition (search term in titles; no further reviews included)
- Obesity (search term in titles; no further reviews included)
- Overweight (search term in titles; no further reviews included)
- “Mass media” (search term in all text; no further reviews included)
- “Choice architecture” (search term in all text; no further reviews included)

Following this, the search was widened beyond research by the Cochrane Collaboration, prioritizing more up-to-date evidence and prioritizing the most cited research. The Google and Google Scholar search criteria were limited to overviews published since 2014 and only the first three pages of results (on each of Google Scholar and Google) were checked; only the first two pages were checked for systematic reviews and meta-analyses. If research items were returned by one of these searches but excluded in the present literature review, the reason was noted in the spreadsheet “[Lists of included and excluded research.](#)” Less systematically, the results for each search term up to page 10 on both Google and Google Scholar were also skimmed for other results that may have been particularly relevant or useful, but exclusions were not noted. For each of the search terms focusing on systematic reviews, the same process was repeated with “AND (behavior OR behaviour OR intervention)” added to the search terms. For those marked with an asterisk, not all exclusions were noted, since the search or a variation of the wording returned very few potentially relevant results that had not already been cited or checked:

- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (diet OR nutrition OR obesity OR overweight)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND food AND (behavior OR behaviour OR intervention)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (smoking OR tobacco)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (alcohol OR binge drinking)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (“physical activity” OR exercise)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (“illicit drug use” OR “drug misuse” OR “substance abuse” OR “substance use”)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (“sexual health” OR “contraception”)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (adherence OR compliance OR concordance) AND (medication OR medicine OR treatment)
- (“overview of systematic reviews” OR “high-level review” OR “meta-review”) AND (adherence OR compliance OR concordance OR attendance OR uptake) AND (screening OR vaccination OR immunization OR test OR testing)
- (“systematic review” OR meta-analysis) AND (diet OR nutrition OR obesity OR overweight)

- (“systematic review” OR meta-analysis) AND food*
- (“systematic review” OR meta-analysis) AND (vegetarian OR vegan OR plant-based OR pescetarian)*
- (“systematic review” OR meta-analysis) AND (meat OR fish OR chicken OR poultry)*
- (“systematic review” OR meta-analysis) AND (meat OR fish OR chicken OR poultry) AND (diet OR nutrition OR obesity OR overweight)*
- (“systematic review” OR meta-analysis) AND (smoking OR tobacco)
- (“systematic review” OR meta-analysis) AND (alcohol OR binge drinking)
- (“systematic review” OR meta-analysis) AND (“physical activity” OR exercise)
- (“systematic review” OR meta-analysis) AND (“illicit drug use” OR “drug misuse” OR “substance abuse” OR “substance use”)
- (“systematic review” OR meta-analysis) AND (“sexual health” OR “contraception”)
- (“systematic review” OR meta-analysis) AND (adherence OR compliance OR concordance) AND (medication OR medicine OR treatment)
- (“systematic review” OR meta-analysis) AND (adherence OR compliance OR concordance OR attendance OR uptake) AND (screening OR vaccination OR immunization OR test OR testing)

Less systematically, additional research items were included based on mentions in included items, follow-up searches for specific interventions, and follow-up searches for reviews that cover multiple behavior types.¹³

In total, 85 out of 166 (51%) of the Cochrane reviews we included and 397 out of 540 (74%) other research items were identified through less systematic methodologies.

Explanation and discussion of the four qualitative analyses used in this literature review

These research items are aggregated in four ways:

- A list of all of the relevant included evidence to particular questions of interest in the health behavior literature. Comment is provided on the strength of evidence for the judgement given, described as “very low,” “low,” “moderate,” “high,” or “very high.” The rating is informed by analysis of the strength of evidence within particular meta-analyses, using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) tool.¹⁴ This analysis is accessible in the “[Strength of Evidence Assessments](#)” spreadsheet.
- Assessment of the likely effect size of particular intervention types and differences in effect sizes across particular variables, as tested directly in meta-analyses that focus on behavioral outcomes. This is accessible in the “[Effect Size Estimates](#)” spreadsheet.

¹³ This increased the risk of selection bias in reporting results but seemed worthwhile in order to seek clarification on topics of particular interest to the farmed animal movement or topics for which the searches so far had returned little evidence.

¹⁴ Holger Schünemann, Jan Brożek, Gordon Guyatt, and Andrew Oxman (eds.), “GRADE Handbook,” last updated October 2013, <https://gdt.gradepro.org/app/handbook/handbook.html>.

- Assessment of the likely differences in effect sizes across particular variables, as tested indirectly in meta-analyses that focus on behavioral outcomes by comparing between different group results.¹⁵ This is accessible in the “[Effect Size Estimates](#)” spreadsheet.
- Analyses of the effectiveness of interventions as moderated by the inclusion of particular behavior change techniques (BCTs), the demographic characteristics of the participants, the intensity of the intervention, the complexity of the intervention, and the use of tailoring. A separate method was used for these moderators because it seemed important to consider how the questions of interest vary across different intervention types, as well as across health behaviors. This is accessible in the “[Moderator Analyses](#)” spreadsheet.

These methods are used in lieu of formal statistical models due to the complex and inconsistent nature of the literature, which makes something like a meta-meta-analysis inappropriate or, at best, cumbersome in that they would require a variety of subjective judgement calls.

“Overall judgement” and “Strength of evidence assessments”

This analysis is a judgement about whether reviews tended to find statistically significant evidence in favor of particular hypotheses (phrased as questions in column B on the “GRADE overview” tab of the “[Strength of Evidence Assessments spreadsheet](#)”) or not.¹⁶

The overall judgements on these questions are essentially a statement of the author of this literature review’s confidence that this sort of intervention is likely to have a meaningful effect in the farmed animal movement, rather than a categorical statement that the intervention will or will not have any effect whatsoever.

The “Strength of evidence” ratings are a subjective judgment of the author, including using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) tool.¹⁷ The GRADE tool is used to clarify understanding of the strength of evidence without spending too long delving into each review. The GRADE scores given were based primarily on the criteria for evidence type (RCTs or observational studies), risks of bias, inconsistency of the results (i.e. unexplained heterogeneity), imprecision (usually, whether the CIs included the null hypothesis), and publication bias. The original reviewers’ GRADE ratings were used when provided, which was most common in Cochrane reviews. Particularly subjective judgments made in this review were marked with an asterisk, highly uncertain judgments (say, roughly less than 75% confident) were

¹⁵ For example, a review might include a meta-analysis for face-to-face counseling interventions compared to control groups that received no intervention and a separate meta-analysis for counseling interventions delivered by telephone, also compared to control groups that received no intervention. Indirect comparison between the reported effect sizes in each of these two meta-analyses might suggest larger effects for one mode of delivery than the other.

¹⁶ For example, one question is whether the evidence in the health behavior literature suggests that education or information only interventions are likely to be effective at changing behavior in the farmed animal movement. A large number of reviews reported statistically significant findings suggesting that such interventions do have effects on behavioral outcome, although several reviews found statistically insignificant evidence. The “overall judgement” is that yes, education or information only interventions are likely to be effective at changing behavior in the farmed animal movement.

¹⁷ Holger Schünemann, Jan Brożek, Gordon Guyatt, and Andrew Oxman (eds.), “GRADE Handbook,” last updated October 2013, <https://gdt.gradepro.org/app/handbook/handbook.html>.

marked with an exclamation mark, and where the review did not provide sufficient information to make a confident judgement, but an answer was guessed, this was marked with a question mark.

A number of other factors influenced the “Strength of evidence” rating, including:

- The number of studies and/or the number of participants,
- The type of study design and a quick impression of the risk of bias for included studies in qualitative reviews,
- The participants in the study, and whether these seemed likely to be broadly representative or represented unique results for specific medical conditions or specific demographic groups,
- The type of comparator used and whether this represented the total effect or something else, such as the difference between different intensities of intervention,
- The outcome measure used, and whether this seemed to be an accurate approximation of health behavior,
- The presence or lack of statistically significant findings in the opposite direction to the main reported outcome,
- The CIs of reported effect sizes,
- The findings of additional analyses within the overview, review, or study that strengthen or undermine the main finding.

A well-designed systematic review and meta-analysis can still provide very weak evidence if the quality of the studies included in the review is poor. For this reason, the AMSTAR 2 ratings of reviews (see the “AMSTAR 2” and “AMSTAR 2 sub questions” tabs on the [spreadsheet](#)) were not particularly informative, and were thus excluded after an initial attempt to include them in this review. The most important elements of the AMSTAR 2 criteria, such as whether a review notes the risk of bias or of publication bias, are in the GRADE criteria anyway.

The “Strength of evidence” ratings have the following meanings, in relation to the “Overall judgement” column:

- Very strong: 85-100% confidence,
- Strong: 75-84% confidence,
- Moderate: 65-74% confidence,
- Weak: 55-64% confidence, and
- Very weak: 50-54% confidence.

These ratings assume that evidence from the health behavior literature constitutes valid evidence for the farmed animal movement, i.e. that there are no issues of replicability between the health behavior literature and the same studies done in the farmed animal movement, which is an inaccurate simplification.

Health behaviors vary in their relevance to the farmed animal movement.¹⁸ The author’s understanding of the relevance of particular health behaviors to the context of the farmed animal movement affects some of the judgement calls in this literature review.¹⁹

When evaluating long-term effects, the focus was on studies with follow-up at 12 months or greater. This distinction was used by a number of included reviews. Usually, if a review included information for multiple follow-up periods but not for follow-up periods of 12 months or greater, it was not considered to provide evidence on long-term effects. Sometimes small exceptions to this were made, such as in one review that provided evidence on follow-up at 10 months or more.²⁰

“Size of effects or differences shown through direct tests”

The effect size that included meta-analyses suggest each intervention type has was also analyzed.

This analysis focused on the effect size estimates from meta-analyses using behavioral outcomes. Clinical or indirect outcomes were not included, although combined measures that featured behavioral outcomes prominently sometimes were. This decision was made because clinical outcome measures and indirect measures such as intention could bias the effect size estimates.

¹⁸ Hence, for example, conclusions from research on interventions seeking to reduce levels of salt consumption seem more likely to also be accurate for interventions seeking to reduce levels of animal product consumption than do the conclusions from research on interventions seeking to encourage vaccination adherence.

Examples of moderating characteristics encountered in this research, of health behaviors and of study subjects, include:

- Addictiveness of a drug or habit,
- Health side effects associated with either the preferred or targeted behavior,
- Frequency and inconvenience of necessary action for the preferred behavior,
- Financial cost,
- Likelihood of long-term failure or recidivism with the preferred behavior,
- Proportion of population,
- Social acceptability of the preferred behavior,
- Enjoyment (or lack thereof) associated with the preferred behavior,
- Awareness of the benefits of the preferred behavior,
- Gender,
- Age,
- Education and socioeconomic status,
- Simplicity or clarity of the ask, and
- Personal or altruistic motivation of the participant.

¹⁹ The “Strength of evidence” estimates in the “GRADE health behavior summaries” tab do not take into consideration how comparable the author expects the evidence from each health behavior to be to evaluations of consumer interventions in the farmed animal movement, but those in the “GRADE overview” tab do. In the “GRADE overview” tab, the evidence from some health behavior areas is weighted more heavily than the evidence from others. For these weightings, intuitions about the relevance of each of the health behavior areas to the farmed animal movement were used.

²⁰ A. Sadiq Sani, Charles Abraham, Sarah Denford, and Susan Ball, “School-based sexual health education interventions to prevent STI/HIV in sub-Saharan Africa: a systematic review and meta-analysis,” *BMC Public Health* 16, no. 1069 (October 2016).

Sometimes the author had to make a subjective judgment call on which reported effect size would be best to include in the analysis; these were chosen based on their intuitive relevance to the farmed animal movement, such as choosing reduction of fat intake over increase in vegetable consumption since reduction may be more important than increasing in that context. Where a meta-analysis included separate analyses for separate age groups, the combined effect size was included where possible, and the effect for adults was included if this was not possible. Where a study reports both dichotomous and continuous outcomes, if both are similarly comparable to behaviors of interest in the farmed animal movement, the continuous outcomes were prioritized, since many in the farmed animal movement are interested in reductions in animal product consumption and not just whether consumption is entirely eliminated, unless the evidence was notably stronger for dichotomous outcomes.

When reporting on the effect sizes of different interventions and on the strength of evidence, the metrics provided by the reviewers were used. As a result, several different measurements of effect size were used, most commonly relative risk (RR), odds ratios (OR), mean difference (MD), standardized mean difference (SMD) / Cohen's d , Hedges' g , and Pearson's r correlation coefficients.²¹ Sometimes the reviewers also report what these effects mean in the original units (such as number of successful quit attempts per 1000 smokers, or the reduction in grams of pure alcohol consumed), but they do not always do so.

Guidelines exist for interpreting some of these different forms of outcome measures, most notably Jacob Cohen's definitions of $r = 0.1$ or $d = 0.2$ as a "small effect size," $r = 0.3$ or $d = 0.5$ as a "medium effect size," and $r = 0.5$ or $d = 0.8$ as a "large effect size." This implies that effect sizes under $r = 0.1$ or $d = 0.2$ are very small (sometimes referred to as "negligible") and that effect sizes well above $r = 0.5$ or $d = 0.8$ are very large.²² In this literature review, Hedges' g is interpreted using the same rules as Cohen's d , despite some differences between the measurements. Because relative risks and odds ratios are not directly comparable to Cohen's d effect sizes, and there are not commonly used guidelines for interpreting them that the author is aware of, this literature review uses the following rough rule of thumb for interpreting them: 0.86 to 1.24 as "very small," 0.85 to 0.76 or 1.25 to 1.49 as "small," 1.50 to 1.99 or 0.75 to 0.51 as "moderate," 2.00 to 2.99 or 0.5 or 0.34 as "large," and 3.00 upwards or 0.33 downwards as "very large."²³ When discussing differences in effectiveness, the same guides are used for interpretation, but the language is sometimes altered to "very slightly different," "slightly different," "moderately different," "substantially different," or "very substantially different."

In the tab "Direct tests health behavior summaries," a mean effect size is provided for each different effect size unit, per question, per health behavior, as is a comment on the range of possibilities indicated by the CIs. Here, the terms "positive" and "negative" are used as a shorthand for "in the hypothesized direction" and "in

²¹ Mark Simmonds, "Synthesis and meta-analysis," (November 2018), <https://drive.google.com/file/d/121LQMSgbpJEvxEJHpUG4wq3KdDhJTEFd/view> provides a clear introduction to understanding most of the main measures used in this literature review.

²² Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences* (New Jersey: Lawrence Erlbaum Associates, 1988; first published 1969), 79-81.

²³ OR and RR are not the same, but to prevent extensive time being spent on trying to locate sufficient information to convert between these measurement types, the same rough rule of thumb was used, with some discretionary adjustment towards 1 for OR.

the opposite direction to the hypothesized direction.”²⁴ For SMD or similar outcome measures, negative SMDs were converted to positive SMDs, if the effect is in the hypothesized direction, to make them more comparable.²⁵ The same was not done the same for RRs and ORs, because they are harder to convert.

When choosing how to summarize the “Overall effect size judgement” in contexts in which there are multiple different units used for the effect sizes, considerations similar to those listed in the section above on “Strength of evidence assessments” were used to make subjective judgement calls.

In the “Direct tests overview” tab, the methodology is similar. Here, the mean effect sizes of each health behavior summary is reported, rather than the mean of each individual effect size. This avoids over-representing health behaviors for which there is more evidence and underrepresenting health behaviors for which there is less evidence. The main disadvantage of this decision is that health behaviors with better evidence on a particular intervention type are weighted equally to those that have only a single effect size.²⁶ This means that sometimes outlier effect sizes are overrepresented. Because most combined effect size estimates are “very small” or “small,” where an effect size estimate approaches “moderate” or higher, it has been highlighted in green. Negative effect sizes are highlighted in red.

There are other disadvantages and limitations of the methodology used in these effect size estimates from direct tests:

- The findings are especially vulnerable to selection bias, because reviews often had a large variety in reported effects. For example, sometimes an indirect comparison in one review would suggest a large effect in one direction, but then another, seemingly similar review would suggest a very small difference in the opposite direction. Decisions about which reviews to include or exclude from the effect size analysis can therefore make a large difference to the estimates, especially since most estimates are based on the findings from a small number of reviews (usually fewer than five).
- Interventions to address smoking may be overrepresented compared to other health behaviors because reviews focused on smoking are more likely to use behavioral outcomes and because the evidence base for smoking is stronger than for some other health behaviors. For example, 25 of the 39 questions for which an effect size was estimated from direct tests included some evidence from smoking, compared to only 11 that included some evidence from diet and nutrition. Other health behaviors are represented to varying extents.
- The width of the CIs is not factored into the mean effect size, although this has partly been addressed by noting the width of included CIs.

²⁴ For example, when asking whether the evidence suggests that gain-framed messages are likely to be more effective at changing behavior in the farmed animal movement than loss-framed messages, the evidence for reviews of multiple health behaviors is summarized as having “Very small negative to small positive” CIs. This means very small differences in favor of loss-framed messages to small differences in favor of gain-framed messages.

²⁵ That is, even if a review reported a reduction in alcohol consumption from an intervention by SMD -0.20, in the spreadsheet, this would have been recorded as SMD 0.20. Any negative numbers on the spreadsheet therefore represent changes in the opposite direction to that which was expected.

²⁶ For example, there is a large amount of evidence on the effect sizes for brief interventions to reduce alcohol consumption, but alcohol consumption BIs are not notably more comparable to plausible interventions in the farmed animal movement than BIs to reduce smoking. However, this does mean that 9 effect sizes for alcohol BIs have the same combined weight as the single effect size for sexual health BIs.

- Scientific guidelines for interpretation of effect sizes are still quite arbitrary, and there is even less agreement on guidelines for interpreting RR and OR.

Although not technically direct tests, beta coefficients from meta-regression analyses have also been included in this section, rather than in the section on indirect comparisons. This is because they report a single measure (the beta coefficient) rather than two separate effect sizes. Since the author did not initially plan to include an analysis of beta coefficients, it is possible that several relevant results from included reviews were omitted.²⁷

“Size of differences shown through indirect comparisons”

Some questions of interest are not well-represented through direct tests. For example, direct tests are not usually made to check the differences between outcomes at short-term and long-term follow-up. To access some of these questions of interest, the likely differences in effect sizes across particular variables were estimated by comparing between the reported effect sizes for different groups within the same meta-analyses.

The methodology used was similar to that for the direct comparisons, except that all differences are reported as percentages, which are then interpreted with the same rough guidelines used to interpret RR.²⁸ The width of reported possible CIs is typically much larger than in the effect size estimates from direct tests, as it involves comparing the extreme ends of the 95% CIs of two different meta-analytic results.

There are other disadvantages and difficulties of interpretation in these comparisons:

- As with the effect size estimates from direct tests, decisions relating to which meta-analyses to include or exclude can make a substantial difference to the “Overall effect size difference judgement.”
- Differences appear more extreme if both results are closer to no effect than if they are both more strongly positive or negative.²⁹
- Some effect size units used probably make the gap appear larger than others.³⁰
- There is unequal representation of the different health behaviors, as with the effect size estimates from direct tests.
- Indirect comparisons between significant results and non-significant results for SMD or Hedges g (but not for RR or OR) are fairly meaningless and difficult to interpret, since the comparisons produce negative numbers. In these situations, descriptions of the differences in CIs are fairly arbitrary; the author usually refers to such differences as “large” or “very large.”

²⁷ That is, the author may have read a review some time in 2018, and read the relevant information relating to meta-regression, but not written this information down, because it was not relevant to the analyses he was planning to conduct. The author only decided to include an analysis of beta coefficients in March 2019.

²⁸ For example, a MD of 14.91g reduction in ethanol consumption at less than 12 months’ follow-up after an intervention compared to a MD of a 7.46g reduction in ethanol consumed at 12 months’ follow-up or longer is reported as a 50% difference in effect size in favor of short-term follow-up. This is then interpreted as a “large” effect size difference (on the border with a “moderate” effect size difference).

²⁹ For example, SMD 0.2 would be 200% the size of SMD 0.1, but SMD 0.3 would only be 150% the size of SMD 0.2, even though the gap between the effect sizes is the same in each of these comparisons in absolute terms (0.1 difference).

³⁰ For example, The gap between SMD 0.1 and SMD 0.2 is not too dissimilar from the gap between RR 1.12 and RR 1.25 (both a gap from the border of a “small” effect size to a more clearly “very small” effect size), yet the percentage difference is very different (200% and 112% respectively).

For the comparisons between effect sizes at long-term and short-term follow-up, the author usually tried to compare across the measurements closest to the 12-month boundary, which was the boundary used to represent “long-term” in the strength of evidence assessments and comments on statistical significance.³¹ Sometimes exceptions were made, such as if the evidence was much stronger for a different follow-up period that also enabled comparison between short- and long-term follow-up.

In some cases, slightly stricter inclusion criteria were used than were used in the [Strength of Evidence Assessments](#) spreadsheet. For example, when evaluating the difference between face-to-face and non face-to-face modes of delivery, indirect comparisons between mixed modes of delivery and one mode of delivery were not included.

Moderator analyses

For the effects of including particular types of behavior change techniques (BCTs) in interventions, modifying the intensity of interventions, modifying the complexity of interventions, the use or lack of tailoring, and the effectiveness of interventions in groups of participants with different demographic characteristics, it seemed as important to consider how results varied by intervention as how they varied by health behavior. To consider this, the results from reviews were listed in tables that laid out the results.

The tables were then examined to see if reviews came to similar conclusions for particular intervention types. Qualitative “scores” were assigned to different types of review. For example, a qualitative review provides evidence rated half as strong as a meta-analysis (further explanation is provided at the top of each tab on the spreadsheet). These methods allow less scope for nuanced understanding of the implications of each review than do the other methods of analysis used in this literature review.

There are also various limitations in the original reviews in the methods used to note whether these factors did or did not have a moderating effect.³²

For the analysis of results by inclusion or exclusion of certain BCTs, the other three forms of analysis used in this literature review (“Size of effects or differences shown through direct tests,” “Size of differences shown through indirect comparisons,” “Overall judgement” and “Strength of evidence assessments”) seemed unlikely to be helpful, so were not carried out. For intensity, complexity, and tailoring, these three types of analysis were used for considering results by health behavior (as done elsewhere), but not by intervention type.

³¹ For example, for Eleni Mantzari, Florian Vogt, Ian Shemilt, Yinghui Wei, Julian P. T. Higgins, Theresa M. Marteau, “Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis,” *Preventive Medicine* 75 (2015), 75-85, where 6 months’, 6-12 months’, 12-18 months’, and 18+ months’ follow-up were all included, 6-12 and 12-18 months’ follow-up were compared, rather than 6 and 18+.

³² See, for example, the discussion at the start of the section on [“Strengths and weaknesses of the health behavior literature.”](#)

Results and discussion

In total, 706 research items were reviewed. The findings are summarized in Table 1 below.

Table 1: Summary of findings. Green highlighting indicates positive results, red indicates negative results, and no highlighting indicates limited or mixed evidence or small effect size.

Issue	Question	Overall judgement	Strength of evidence	Size of effects or differences shown through direct tests	Size of differences shown through indirect comparisons
Education or information only	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	Small / very small effects	Very small / small differences in favor of the short-term
Brief interventions (BIs)	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	Small effects	Moderate / large differences in favor of the short-term
Motivational interviewing (MI)	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small / very small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	Large differences in favor of the short-term
	Does the evidence suggest that this intervention is likely to be more effective at changing behavior in the farmed animal movement than other comparable intervention types?	Yes	Very weak	Very small / small differences in favor of MI	No information
Counselling or therapy	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Large / moderate effects	NA

	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	Small effects	Moderate differences in favor of the short-term
Self-help, self-monitoring, and self-management	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	No information	No information
Action planning, coping planning, problem solving, and implementation intentions:	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Small / moderate effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
Social norms	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information
Peer-led interventions and mentoring	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Very small / small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
Reminders	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Very small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	No information	No information

Educational or behavioral interventions as additions to pharmacological interventions:	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Unclear (very small?) effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information
Telephone	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	No useful information	No information
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	No	Very weak		Small differences in favor of face-to-face delivery
Text messaging	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	No information		No information	No information
Broader mobile phone interventions (mHealth)	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	No information	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information

	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	No information		No information	No information
Online and computer-based interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small / very small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	Very small effects / no effect	Moderate / large differences in favor of the short-term
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	Yes	Very weak	Very small differences in either direction / no difference	Moderate / large differences in favor of face-to-face delivery
Explicit consideration of print-based materials	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Very small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	Yes	Weak	No information	No information
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than other non face-to-face modes of delivery?	No	Very weak	Very small differences in either direction / no difference	No information
Broader consideration of non face-to-face modes of delivery	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small / very small effects	NA

	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	Very small / no effect	Large / moderate differences in favor of the short-term
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	No	Very weak	No information	Moderate / small differences in favor of face-to-face delivery
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than other non face-to-face modes of delivery?	NA	NA	NA	NA
All non face-to-face modes of delivery combined	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Very strong	Small / very small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	Very small / no effect	Moderate / large differences in favor of the short-term
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?	No	Very weak	Very small differences in either direction / no difference	Small differences in favor of face-to-face delivery
Group interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Moderate effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
	Does the evidence suggest that interventions using group delivery are likely to be less effective at changing behavior in the farmed animal movement than individual equivalents?	No	Moderate	No difference / very small differences in favor of group delivery	Very small differences in favor of group delivery

School-based interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Weak	Small / very small effects?	Very small differences in favor of the long-term
	Does the evidence suggest that interventions in this setting are likely to be more effective at changing behavior in the farmed animal movement than other settings?	No	Very weak	No information	No information
Family-based interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Weak	Small effects	Small / moderate differences in favor of the short-term
	Does the evidence suggest that focusing on parents is important for the success of home- or family-based interventions?	No	Very weak	No information	Large differences in favor of favor of interventions that involve both parents and children
Workplace-based interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	No information	No information
Community-based interventions and population-level interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Very weak	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Very weak	No information	No information
	Does the evidence suggest that interventions in this setting are likely to be more effective at changing behavior in the farmed	No	Very weak	No information	Mixed

	animal movement than other settings?				
Other settings	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Very weak	No information	NA
	Does the evidence suggest that interventions in specific settings are likely to be more effective at changing behavior in the farmed animal movement than other settings?	No	Very weak	No information	No information
Intensity of intervention	Does the evidence suggest that more intense interventions are likely to be more effective at changing behavior in the farmed animal movement?	Yes	Very weak	Small differences in favor of more intensity	Small differences in favor of more intensity?
	Does the evidence suggest that interventions with longer total duration are likely to be more effective at changing behavior in the farmed animal movement?			Very small / small differences in favor of longer duration	Mixed
	Does the evidence suggest that interventions with a larger number of contacts are likely to be more effective at changing behavior in the farmed animal movement?			Small / moderate differences in favor of more contacts	Very small differences in favor of more contacts
	Does the evidence suggest that interventions with longer contact time are likely to be more effective at changing behavior in the farmed animal movement?			Very large differences in favor of longer contact time	Small differences in favor of longer contact time
	Does the evidence suggest that text or digital interventions with two-way communication are likely to be more effective at changing behavior than those with one-way communication?			No information	Very small / small differences in favor of one-way communication
Complexity of intervention	Does the evidence suggest that more complex interventions are likely to be more effective at changing behavior in the farmed animal movement?	Yes	Very weak	Small differences in favor of more complexity	Moderate / large differences in favor of more complexity

	Does the evidence suggest that interventions with more components (BCTs or similar) are likely to be more effective at changing behavior in the farmed animal movement?			Small differences in favor of more components	Moderate / large differences in favor of more components
	Does the evidence suggest that interventions with more modes of delivery are likely to be more effective at changing behavior in the farmed animal movement?			No information	No difference / very small differences in favor of more modes of delivery
	Does the evidence suggest that interventions with behavioral components are likely to be more effective at changing behavior in the farmed animal movement than those only using education?			Small differences in favor of more behavioral components	Large / moderate differences in favor of more behavioral components
Tailoring	Does the evidence suggest that tailored interventions are likely to be more effective at changing behavior in the farmed animal movement than comparable non-tailored interventions?	Yes	Very weak	Very small / small differences in favor of tailoring	Moderate / large differences in favor of tailoring
Fear appeals	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Very weak	Small / very small effects	NA
	Does the evidence suggest that fear appeals are likely to be more effective at changing behavior in the farmed animal movement than comparable non-fear-based interventions?	Yes	Very weak	No information	No information
Framing	Does the evidence suggest that gain-framed messages are likely to be more effective at changing behavior in the farmed animal movement than loss-framed messages?	No	Very weak	Very small differences in favor of gain-framed messages	Large differences in favor of gain-framed messages?
	Does the evidence suggest that positive attribute framed messages are likely to be more effective at changing behavior in the farmed	No	Very weak	No information	No information

	animal movement than negative attribute framed messages?				
Narratives	Does the evidence suggest that testimonial messages are likely to be more effective at changing behavior in the farmed animal movement than non-testimonial messages?	No	Very weak	No information	No information
	Does the evidence suggest that narrative messages are likely to be more effective at changing behavior in the farmed animal movement than non-narrative messages?	Yes	Very weak	Small differences in favor of narrative messages	No information
Cultural competency	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that culturally competent interventions are likely to be more effective at changing behavior in the farmed animal movement than generic interventions?	No	Very weak	Very small differences in favor of interventions that are not culturally competent	No information
Variations across professionals	Does the evidence suggest that professionals with more specialized skillsets and expertise are likely to be more effective at encouraging or supporting behavior change than less specialized professionals?	Yes	Very weak	No information	Unclear
	Does the evidence suggest that professionals with longer training times and greater public respect and credibility are likely to be more effective at encouraging or supporting behavior change than those with less respect and credibility?	No	Weak	No information	Unclear
	Does the evidence suggest that interventions provided by health professionals in other contexts (such as oral health professionals or pharmacists) are likely to be effective?	Yes	Moderate	Moderate / large effects	No information

Other variations	Does the evidence suggest that interventions where the patient and deliverer are matched on demographic characteristics are likely to be more effective at changing behavior in the farmed animal movement than interventions without demographic matching?	Yes	Weak	No information	Small differences in favor of demographic matching
Pharmacology	Does the level of behavioral support improve the effectiveness of pharmacological treatments?	No	Weak	No information	Very small differences in favor of lower intensity behavioral support
	Is motivation an important moderator of the effectiveness of pharmacological treatments?	Yes	Very weak	No information	No information
	Do those who use this intervention to assist behavioral change rely on continued access to it in order to maintain their new behavior?	Yes	Very weak	No information	No information
Convincing alternatives	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Large effects?	NA
	Does the evidence suggest that e-cigarettes (an alternative to smoking) are similarly effective to other forms of NRT?	Yes	Very weak	No information	No information
	Does the evidence suggest that programs that focus on sexual abstinence are less effective than those that encourage the use of acceptable alternative methods of sexual intercourse?	Yes	Weak	Small differences in favor of alternative methods	Moderate differences in favor of alternative methods
Direct financial incentives	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Small / moderate effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No	Weak	Moderate effects	Small differences in favor of the short-term
	Do those who use this intervention to assist behavioral change tend to rely on continued access to it in	Yes	Weak	Very small effects after withdrawal of the incentives	Moderate differences in favor of continued use of incentives

	order to maintain their new behavior?				
	Does the evidence suggest that cash incentives are superior to other forms of incentive?	Yes	Very weak	No information	Small differences in favor of cash incentives
	Does the evidence suggest that higher value incentives are likely to have a larger impact on behavioral outcomes of interest to the farmed animal movement than lower value incentives would?	Yes	Weak	Very small differences in favor of higher value incentives / no difference	Small differences in favor of higher value incentives
	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement, even if the recipients of the intervention are receiving other types of intervention simultaneously?	No	Very weak	No effect / very small effects beyond other interventions	Moderate / small differences in favor of inclusion of this intervention
Taxes, subsidies, and prices	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Strong	Moderate effects?	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information
	Does the evidence suggest that subsidies for animal-free food products are likely to have a larger impact on the consumption of those products than taxes would on animal products?	Yes	Very weak	No information	Large differences in favor of subsidies
Packaging and labelling	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Very small effects?	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
	Does the evidence suggest that pictorial or interpretive labels are likely to be more effective at changing behavior in the farmed animal movement than text-only labels?	Yes	Weak	No information	Large differences in favor of pictorial or interpretive labels

	Does the evidence suggest that graphic labels are likely to be more effective at changing behavior in the farmed animal movement than non-graphic labels?	Yes	Weak	No information	No information
Nudges and environmental interventions	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Moderate	Small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
Mass media	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Very small / small effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	No information		No information	No information
	Does the evidence suggest that more complex mass media interventions that use multiple intervention components are likely to be more effective than purely mass media interventions that utilize equivalent amounts of resources?	Yes	Very weak	No difference / very small differences in favor of complexity	No information
	Does the evidence suggest that more intensive mass media interventions are likely to be more effective than less intensive mass media interventions?	Yes	Very weak	Small differences in favor of intensity	No information
	Does the evidence suggest that mass media alone is likely to be less effective at changing behavior in the farmed animal movement than mass media combined with face-to-face engagement?	Yes	Very weak	Small differences in favor of combination with face-to-face	No information
Social marketing	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Small / moderate effects?	NA

	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
Month-long campaigns	Does the evidence suggest that this intervention is likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	No information	NA
	Do those who use this intervention to assist behavioral change will maintain their new behavior after the end or withdrawal of the intervention?	Yes	Very weak	No information	No information
Advertising and advertising bans	Does the evidence suggest that the advertising of products related to health behavior increases consumption of them?	Yes	Weak	Small / very small effects	No information
	Does the evidence suggest that bans of advertisements for undesirable products are likely to be effective at changing behavior in the farmed animal movement?	No	Very weak	No effect / very small effects	No information
Bans on the risk behavior	Are localized bans of risk behaviors effective at reducing those risk behaviors in those contexts?	Yes	Weak	Moderate / small effects	NA
	Does the evidence suggest that where these localized effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
	Are localized bans of risk behaviors effective at reducing total risk behaviors, beyond where the ban is operational?	Yes	Weak	No information	NA
Other legislation	Does the evidence suggest that these interventions are likely to be effective at changing behavior in the farmed animal movement?	Yes	Weak	Moderate effects	NA
	Does the evidence suggest that where effects occur, they are likely to endure into the long term?	Yes	Very weak	No information	No information
Gradual or abrupt transition	Does the evidence suggest that gradual behavior change is likely to be more successful than abrupt	No	Weak	No information	No information

	behavior change in the farmed animal movement?				
Combined average across all individual included results (as opposed to combining by health behavior)	Does the evidence suggest that results at long-term follow-up are smaller than results at short-term follow-up?				Small / moderate differences in favor of the short-term
	Does the evidence suggest that interventions using this mode of delivery are likely to be less effective at changing behavior in the farmed animal movement than face-to-face equivalents?				Small differences in favor of face-to-face?

Strengths and weaknesses of the health behavior literature

The strengths and weaknesses of the health behavior literature affect the weight that should be placed on the findings of this literature review. Each of the following comments and generalizations is based on the author's reading of the included evidence, primarily through systematic reviews and overviews. It is possible that individual studies contain much more information on some of these issues than the author is aware of and that such information simply has not been included in the identified research items.

Strengths of the health behavior literature

- There is a far larger quantity of research on health behavior than there is (or likely ever will be) on behavior change that affects farmed animals.³³
- The health behavior research covers a greater variety of intervention types and questions than research in the farmed animal movement does. For example, health behavior researchers have devoted significant resources to developing theories of human behavior change, which are used to inform intervention prioritization and design (see [Appendix B: Theory Used in Health Behavior Intervention Design and Research](#)).
- A variety of methodologies have been developed for assessing the risk of bias in individual studies, quality of systematic reviews, and quality of evidence overall.³⁴

³³ Some reviews include over one hundred studies on the same intervention type, whereas the largest meta-analysis of studies within the farmed animal movement that the author is aware of ("Leafleting," Animal Charity Evaluators (November 2017), <https://animalcharityevaluators.org/advocacy-interventions/interventions/leafleting/#report>) contains only six studies.

³⁴ This literature review has sought to utilize tools such as AMSTAR 2 and GRADE, rather than to evaluate them in comparison to other tools and therefore cannot provide strong recommendations on the best evaluation tools. However, most of these could be used by researchers in the farmed animal movement; either as checklists to ensure the quality of their own research designs, or to evaluate the quality of the research of others. Some tools, such as the Cochrane Risk of Bias tool (Julian P. T. Higgins and Sally Green (eds.) "Chapter 8: Assessing risk of bias in included studies" in "Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0" (March 2011), <http://handbook-5-1.cochrane.org/>) can be used with minimal adaptation. Although the full methods used in health behavior systematic reviews might only

Weaknesses of the health behavior literature

- Many reviews use qualitative reporting, rather than meta-analyses. Sometimes, this is due to inconsistency in research design making meta-analysis inappropriate.
- In general, the health behavior has less consistent and strong conclusions from research than expected. Most qualitative reviews of specific intervention types for specific outcomes include a mixture of studies finding statistically significant and insignificant effects in the main outcomes of interest. Sometimes they also include studies finding significant effects in the opposite direction to the hypothesized direction. Meta-analyses often contain substantial heterogeneity that moderator analyses are unable to explain.³⁵
- For most topics, there is a lack of studies with follow-up data collection at a year or more post-intervention, meaning that long-term effects are not well understood.³⁶
- Often review authors do not discuss in detail which characteristics might explain the heterogeneity in the results of research items.
- Generally there is not a strong understanding of the specific components of interventions that makes them more or less effective. In the last decade or so, the method of checking the association between BCT use and intervention effectiveness has been developed. However, due in part to the tendency for interventions to use multiple BCTs simultaneously, there is a lack of experimental evidence comparing near identical interventions in order to better understand the benefits of adding or removing a single BCT.
- There is surprisingly little experimental evidence about the effect of variations in framing, messaging, and communication on behavioral outcomes.³⁷
- Many reviews fail to follow the quality recommendations of review guidelines and appraisal tools, such as AMSTAR 2.³⁸

rarely be appropriate for the research of the farmed animal movement, reading such reviews has helped the author to reflect on the importance of being as systematic as practically possible (given resource constraints) in research in order to minimize selection bias. The author has tried to utilize some of these methods in conducting this literature review, for example (see the section on “search strategy” above).

³⁵ See the column for “Inconsistency of results” in the tab “GRADE individual reviews” on the spreadsheet “[Strength of Evidence Assessments](#).”

³⁶ Many reviews draw attention to this limitation. This literature review includes comparisons between different follow-up periods where possible. See the column for “Question (abbreviated for clarity)” in the tab “GRADE individual reviews” on the spreadsheet for “[Strength of Evidence Assessments](#),” note the much larger number of reviews helping to answer questions marked as “Effective?” than helping to answer questions marked as “Long-term?”

³⁷ For example, in the debate about the usefulness of fear appeals, proponents have used mostly observational evidence, or experimental evidence with indirect outcome measurements such as intention to smoke or self-efficacy, to argue that fear appeals are effective (see, for example, David Hammond, “Health warning messages on tobacco products: a review,” *Tobacco Control* 20 (2011), 327-37). Others have drawn attention to the limitations of such evidence and pointed to indirect evidence of negative effects (see, for example, Robert A. C. Ruiter, Loes T. E. Kessels, Gjalte-Jorn Y. Peters, and Gerjo Kok, “Sixty years of fear appeal research: Current state of the evidence,” *International Journal of Psychology* 49, no. 2 (April 2014), 63-70 and Gerjo Kok, Gjalte-Jorn Ygram Peters, Loes T. E. Kessels, Gill A. ten Hoor, and Robert A. C. Ruiter, “Ignoring Theory and Misinterpreting Evidence: The False Belief in Fear Appeals,” *Health Psychology Review* 12, no. 2 (December 2017), 111-25).

³⁸ The author of this literature review did not count one domain considered to be “critical” by the authors of the AMSTAR 2 methodology as critical, partially because it did not seem particularly important, and partially so that he could meaningfully differentiate between reviews; otherwise nearly all evaluated reviews would have been rated as “low”

- Often, individual studies have weaknesses in their design and risks of bias. In some reviews — or entire areas of research — the vast majority of studies share similar risks of bias or methodological flaws.³⁹ Sometimes, this is for practical and understandable reasons, such as blinding of participants or assessors being difficult for some intervention types. In other cases it is avoidable, such as where studies use sample sizes that are too small and therefore insufficiently powered to detect significant differences. This latter methodological weakness can make qualitative reporting of study results especially misleading.⁴⁰
- Some methods of reporting qualitative results seem clearer and more transparent, such as the harvest plot method.⁴¹ However, this method was rarely used.
- Although social cognition models have been developed and used in research, there is not consistent evidence that any of them predict behavioral outcomes more accurately than the others, or that the use of particular theories leads to greater intervention effectiveness.⁴²
- In some instances, there are surprisingly few comparisons between an intervention and no intervention or a placebo. Comparison is frequently made to “usual care” or other active treatment types; the presence of other forms of intervention may reduce the effect size of the intervention being tested. Sometimes the studies evaluated in meta-analyses or reviews include a mixture of different types of comparators.
- A large proportion of included studies were conducted in developed countries, most frequently the US.⁴³ This reduces the ability of the health behavior literature to compare findings across countries.

or “critically low.” No reviews were rated as “high” quality by AMSTAR 2 criteria. See the tab “AMSTAR 2” on the spreadsheet for “[Strength of Evidence Assessments](#).”

³⁹ For example, Denise de Ridder, Floor Kroese, Catherine Evers, Marieke Adriaanse, and Marleen Gillebaart, “Healthy diet: Health impact, prevalence, correlates, and interventions,” *Psychology & Health* 32 (April 2017), 23-4, referring to four qualitative reviews on nudge interventions, note that, “[a]ll reviews stress that the quality of studies included was suboptimal, and in some cases even weak... although nudging shows promise when considering health behaviour interventions, as a low-cost, effective way of increasing healthy choices, caution is needed due to the lack of population-based long-term effectiveness.” Barnabas Szasz, Anna Palinkas, Bence Palfi, Aba Szollosi, and Balazs Aczel, “A Systematic Scoping Review of the Choice Architecture Movement: Toward Understanding When and Why Nudges Work,” *Behavioral Decision Making* 31, no. 3 (July 2018), 355-66 review 422 choice architecture interventions in 156 studies note that, “only 7% of the studies applied power analysis, 2% used guidelines aiming to improve the quality of reporting, no study in our database was preregistered, and the used intervention nomenclatures were non-exhaustive and often have overlapping categories.”

⁴⁰ For example, a review might note that the majority of studies found no significant effects of the intervention, but if this was due to the individual studies being insufficiently powered, then the intervention may still be effective. Even with small or very small effect sizes, low intensity and low cost interventions may still be cost-effective. This problem is mitigated by using meta-analytic techniques.

⁴¹ See, for example, Rory McGill, Elspeth Anwar, Lois Orton, Helen Bromley, Ffion Lloyd-Williams, Martin O’Flaherty, David Taylor-Robinson, Maria Guzman-Castillo, Duncan Gillespie, Patricia Moreira, Kirk Allen, Lirije Hyseni, Nicola Calder, Mark Petticrew, Martin White, Margaret Whitehead, and Simon Capewell, “Are interventions to promote healthy eating equally effective for all? Systematic review of socioeconomic inequalities in impact,” *BMC Public Health* 15, no. 457 (May 2015).

⁴² See the section on “Usefulness of social cognition models” in [Appendix B](#).

⁴³ For example, Dao Truong, “Social Marketing: A Systematic Review of Research 1998-2012,” *Social Marketing Quarterly* 20, no. 1 (January 2014), 15-34, though not limited exclusively to health behavior (about 75% of identified articles focused on health topics), notes that 1,879 (56.9% of those identified) research contributions to the topic of social marketing came from the US. Following this were the UK (11.5%), Australia (5.7%), and Canada (5.1%).

Limitations

- There is a general risk of selection bias through using methods that were to some extent systematic but also relied on arbitrary or intuitive constraints and heuristics for when further effort to make the searches more comprehensive and thorough would lead to steeply diminishing returns.
- As an attempt to balance being systematic and efficient, Google and Google Scholar were used, rather than comprehensively searching through scientific databases, and some fairly arbitrary search terms were used. Some mistakes were made in the selection of search terms which may have either resulted in a greater focus on less relevant findings (such as the initial searches made before adding the terms “AND (behavior OR behaviour OR intervention)”), or which arbitrarily included some terms over others (such as when the word “intervention” was included but “interventions” was not).
- Although the methods used for identifying research items of potential interest were systematic to some extent, some relevant information from the identified research items may have been omitted. At the beginning of this project, it was unclear which topics would eventually be discussed in depth, and which forms of analysis would be carried out. As a result, some data that would otherwise have been relevant to an analysis included within this review may have been omitted, especially earlier on in the process.⁴⁴ Additionally, in the interest of covering substantial amounts of the vast health behavior literature without spending excessive amounts of time on the project, some research items were read quickly, or only certain sections were read. For this reason, some cited research items may have contained additional information of interest which is not reported in this literature review.
- There is a risk of overlap between the studies in included research items. If, after brief inspection, an identified review or overview seemed to have substantial overlap with those that had already been examined, this was noted in footnotes. However, it did not seem worthwhile to check for overlap of individual studies, except in cases where points were made on the basis of a small number of studies.⁴⁵
- By focusing on reviews, rather than individual studies, some relevant and useful information may have been missed from individual studies. Similarly, where overviews were used, the cited reviews were not always examined.

Although this literature review has not recorded or counted the geographical locations of included research items, the author would guess that over half of the primary studies included in this literature review (including those summarized by systematic reviews) were conducted in the U.S. Additionally, the systematic searches did not return a single review that directly compared results between North American and European contexts for particular interventions, although some specified that their review was limited to high-income countries or to low- and middle-income countries only. This problem may be partially due to language constraints. Many reviews only include studies published in English.

⁴⁴ For example, the evaluation of the findings from behavior change technique analyses in included reviews was not initially planned, so information on BCT analyses from reviews that were looked at before making this decision may have been omitted. The analysis of “Effectiveness of interventions varying by demographic characteristics” suffers from the same issue.

⁴⁵ If an overview contained reviews that had already been examined, this was noted in the tab on “Qualitative comments on overviews” in the spreadsheet “[Strength of Evidence Assessments](#).” It is likely that there are many instances where cited reviews use some of the same studies as their evidence base and this has not been noted.

- By attempting to aggregate information from studies in vastly different contexts and using varying methodologies, the findings may present a misleading impression of similarity between the findings.⁴⁶
- By focusing on behavioral outcomes, lots of potentially useful evidence on indirect outcomes (e.g. intentions, knowledge, and self-efficacy) has been excluded.
- After an initial brief exploration, little attention was paid to the role of theory in intervention design.
- The divisions into categories of intervention type are arbitrary and are based on divisions that seemed most useful and relevant to the farmed animal movement. Minor shifts in categorization could lead to different “results” in terms of any of the four forms of qualitative analysis undertaken in this literature review.
- Most health behavior areas can involve interventions focused on either reducing health risks (often referred to as “preventative” interventions) or ceasing unhealthy behaviors. This review does not consistently distinguish between prevention and cessation interventions.
- Although systematic reviews reduce the effects of selection bias, they do little to mitigate other risks of bias (including publication bias) in the included studies. Sometimes sensitivity analyses and trim-and-fill procedures are used to account for these risks, but they are only used infrequently. When dealing with small and very small effect sizes (as is the case with most interventions evaluated in this review), the risk of publication bias seems especially concerning.
- Various analyses in this literature review are based on comparisons across arbitrary cutoff points and dichotomies.⁴⁷
- The author’s lack of expertise in statistics may have meant that unusual forms of statistical analysis in the included research items (that could represent manipulation or reframing of the results) were not noticed.
- Little attention was paid to whether a meta-analysis used fixed or random effects and the degree to which this may have affected results.
- There are weaknesses to each of the forms of qualitative analyses used, as described above.⁴⁸ The use of more established statistical methods would have been very time consuming and would have had other drawbacks, but may have produced more informative and reliable results.
- In the analysis used in this literature review, relevance to the farmed animal movement was generally prioritized over consistency in methodology.
- The health behavior literature has a variety of methodological weaknesses itself, and so this review will likely reflect those limitations.⁴⁹

⁴⁶ For example, different forms of outcome measurement have hardly been considered, but it may well be that studies using self-reported and objectively measured outcomes give substantially different impressions of the effectiveness of particular interventions or modifications. Little attention has been paid to the countries in which studies were conducted; the study locations or characteristics have not been noted unless it seemed to have an especially important bearing on the results.

⁴⁷ Sometimes, this reflects decision-making by the author of this literature review, as with the comparison of “long term” (follow-up at 12 months or more) to “short term” (follow-up at less than 12 months). At other times, this reflects the decisions of the authors of included research items. For example, sometimes intensity is measured by texts being sent weekly or not; alternatively, one could reasonably use texts sent “at least once every five days” as the cutoff point to compare across.

⁴⁸ See the section on [“Explanation and discussion of the four qualitative analyses used in this literature review.”](#)

⁴⁹ See the section on [“Strengths and weaknesses of the health behavior literature.”](#)

- Arguably, understanding the predictors of effects within particular categories of intervention is more useful than understanding the effect size estimates for broad categories of intervention types. This literature review has focused on the latter more than the former.
- Other academic topics that overlap with the health behavior literature that may also be of some interest to the farmed animal movement, such as advertising, have not been included in this literature review.
- This literature review focused on the strategic implications that the health behavior literature has for dietary change interventions in the farmed animal movement. However, dietary change is not the only form of behavior change of interest to the farmed animal movement; activism, voting, and donations might all also determine outcomes for farmed animals, for example.⁵⁰ Additionally, farmed animals are not the only sentient beings that are currently excluded from humanity's moral circle.⁵¹

Potential Items for Further Study

Further research could be conducted to address any of the above limitations of this literature review. However, such research seems unlikely to be high-priority; this literature review was extremely time-consuming and had fewer strategic implications for the farmed animal movement than we had hoped it might.

More specific, higher-priority potential items for further study include:

- Explicit testing of the effects of incorporating or excluding techniques and intervention components identified as promising by this literature review, such as goal setting, social support, and action planning, in the specific contexts that social movements are interested in.⁵²
- Testing for possible interactions between different interventions that enhance their effectiveness. For example, participants could be surveyed for their prior awareness of related issues, to see if this has an interaction effect on the success of the interventions.⁵³
- More formal “overviews” of reviews focusing on specific outcomes of interest.⁵⁴

⁵⁰ See “Asking individuals to contribute to institutional change” in Jamie Harris, “Which institutional tactics can animal advocates use?” (April 27, 2020), <https://www.sentienceinstitute.org/blog/institutional-tactics>.

⁵¹ For example, wild animals, the global poor, people of color, women, members of the LGBTQI community, future beings, and artificial sentience are all arguably excluded or insufficiently included within humanity's moral circle.

⁵² Though lower priority, other variations evaluated by health behavior research, where the conclusions are unclear, could also be evaluated in the contexts that social movements are interested in. For example, what effects do the use of gain-framed messaging or fear appeals have?

⁵³ A similar point was made by Filippo Bianchi, Claudia Dorsel, Emma Garnett, Paul Aveyard, and Susan A. Jebb, “Interventions targeting conscious determinants of human behaviour to reduce the demand for meat: a systematic review with qualitative comparative analysis,” *International Journal of Behavioral Nutrition and Physical Activity* 15, no. 102 (October 2018). They note that, “while providing information was not found to directly influence behaviour, future research should explore whether this approach might contribute towards reducing population-wide demand for meat in other ways. For example, providing information on the benefits of eating less meat might increase the public's acceptability for more structural interventions to reduce meat consumption.”

⁵⁴ These could include more systematic inclusion of relevant reviews and studies, greater effort to dig down into the reasons for discrepancies between reviews on specific outcomes of interest, and better use of meta-analytic techniques.

- A comprehensive, systematic review of studies of the determinants of animal product consumption. One review has done this already from an environmentalist's perspective⁵⁵ and a subsequent masters' thesis included a review of some of the literature.⁵⁶
- A more thorough review focused specifically on interventions for dietary change, seeking especially to identify estimates of cost-effectiveness.
- A detailed look at the methods of study design and measurement. For example, how have studies of dietary change (and reductions in animal product consumption specifically) measured intervention effectiveness? What are the advantages and disadvantages of different approaches?
- A review of relevant aspects of the advertising and marketing literature.⁵⁷
- Reviews of the psychological, business, and educational literature more widely on specific questions of interest, such as on goal setting.

Bibliography

See the [Lists of Included and Excluded Research](#) for a full list of references.

⁵⁵ Susanne Stoll-Kleemann and Uta Johanna Schmidt, "Reducing meat consumption in developed and transition countries to counter climate change and biodiversity loss: a review of influence factors," *Regional Environmental Change* 17, no. 5 (June 2017), 1261-77.

⁵⁶ Joeri Veul, "Interventions to reduce meat consumption in OECD countries: a systematic review to understand differences in success," (master's thesis, August 2018), https://theses.ubn.ru.nl/bitstream/handle/123456789/6391/Veul%2c_Joeri_1.pdf?sequence=1.

⁵⁷ Some of the included research items that had relevance for messaging considerations drew on this literature. Topics included comparison of loss-framing to gain-framing, help-self appeals compared to help-others appeals, moderators of effectiveness in mass media campaigns and in social marketing, and evaluations of the effectiveness and moderators of advertising and advertising bans. This literature may be more useful for improving messaging strategies and identifying the features of mass media campaigns that increase their effectiveness than the health behavior literature is.