and may not need extensive piloting when transferred to new settings.[8] However, opting for immediate implementation of an intervention which is not fit for local purpose would also waste time and money, and could harm participants directly or by displacing more effective interventions.[9] For example, after-school activities aiming to promote positive youth development to prevent teenage pregnancy have been found to be effective in some settings, ineffective in others and harmful in some.[10-12]

The key concept is 'equipoise', usually defined as uncertainty regarding benefits.[13] In the context of a previously evaluated intervention being considered for implementation in a new setting, equipoise refers to remaining uncertainty about whether or not the intervention will be more effective than usual treatment in the new setting. This will partly depend on what is usual treatment in the new setting. But uncertainty will also depend on whether the intervention can be feasibly and acceptably delivered, reach potential beneficiaries and generate outcomes in the new setting. The picture is complicated by the fact that local decision-makers will generally want to know about the size of effects and the costs associated with the intervention to assess whether it is worth the money. Thus, judgements about equipoise are neither simple (they involve multiple factors) nor easy (they are questions of probability not binary distinctions), and cannot be reduced to decisions based on effects demonstrated in previous evaluation.

Realist approach to local decisions about implementing interventions evaluated as effective elsewhere

Let us consider how local decision-makers might decide whether to proceed to full implementation or not of a previously evaluated intervention. We use the example of the proposed implementation in US elementary schools of a school-based social and emotional learning curriculum, which has been previously evaluated as effective in promoting mental health across various settings.

We take a realist approach. Realists view the social world as comprising observable events generated by unobservable mechanisms.[3 14] Realist evaluators focus not merely on what works, but on what works for whom and under what conditions. Interventions are understood to provide resources which actors and recipients use. This use might then trigger a variety of intervention mechanisms which may or may not generate outcomes, intended or otherwise. The use of intervention resources, the mechanisms thus triggered and the extent to which mechanisms generate outcomes will vary between contexts.

For example, it might be the case that social and emotional skills curricula are more feasible to deliver in some schools and with some students more than others. The mechanisms that such interventions trigger – students understanding and managing their emotions and relationships with others - might be sufficient to generate significant mental health benefits in some settings but not others. They might improve mental health where deficits in these personal skills are an important part of the 'aetiological mechanisms' underlying poor mental health, but not in settings where broader structural factors, such as very high rates of violence, are more important aetiological mechanisms. By 'aetiological mechanisms', we refer to the mechanisms that generate certain outcomes in a setting in the absence of intervention, in contrast to 'intervention mechanisms', which refer to the mechanisms triggered by intervention resources which aim to disrupt, over-ride or compensate for aetiological mechanisms to change the pattern of outcomes in a setting.

Below, we discuss several scenarios concerning implementation and effectiveness to consider where local implementation of a previously evaluated intervention is most likely to be successful. For both implementation and effectiveness, we first ask if the intervention has been evaluated in contexts like the new context being considered, and then we ask if the current context has the specific features required for the intervention to be successful.

Low uncertainty scenario

Implementation

In realist terms, implementation will be most likely to be feasible and acceptable and reach potential beneficiaries in a new setting when two conditions are met (figure 1). Firstly, there is evidence from existing process evaluation(s) conducting in settings and to populations which resemble those present in the new setting that the intervention was feasible and acceptable to deliver and achieved reach. Ideally, evaluations will also provide evidence about what factors - related to providers, contexts and target populations - supported implementation. Evaluations can usefully document what was needed to support feasibility in the context of the evaluation using existing frameworks.[15] In the case of social and emotional learning, for example, there is evidence from existing process evaluations that implementation is feasible and acceptable and achieves reach, and evidence that implementation is enabled by the involvement of school leaders championing the intervention, the availability of training, and broad ownership of the intervention among school staff.[16 17]

Secondly, successful implementation in the new setting will be more likely when local information suggests that key factors enabling implementation are also present in the new setting. As well as specific enabling factors identified in previous evaluations of the intervention in question, the most important enabling factors can also be identified from generic implementation frameworks.[2 18]. The general theory of implementation is particularly useful here, suggesting that successful implementation is promoted through the following factors: 'potential' (providers possessing the desire and ability needed for delivery); 'capability' (the intervention being workable and possible to integrate into existing routines, policies and systems); 'contribution' (providers being able to work together to make sense of the intervention, commit to its delivery, coordinate their actions and

reflect on progress); and 'capacity' (local structures, norms, roles and material resources supporting implementation). Local information on whether such factors are present locally might come from: consultations or surveys with local organisations; audits of knowledge, skills, equipment and other resources; existing guidelines for practice; audits or evaluations of local practice; and observations of local service contexts. Although some of this information will probably already exist, some might need to be newly collected by local decision-makers as part of needs assessment for the new intervention.

Effectiveness

Assuming that implementation is possible, whether this will translate into effectiveness will depend on two factors. Firstly, it is of course more likely when existing evaluations suggest the intervention is effective. The evidence base for an intervention should ideally draw on multiple, rigorous studies so that we have evidence of replicability and more precise estimations of effects, ideally from pragmatic evaluations conducted in relevant, real-world conditions. Evaluations will ideally also indicate how the intervention works and for whom. Within studies, light might be shed on this by mediation analyses (what mechanisms are triggered and how might these disrupt aetiological mechanisms underlying outcomes?) or moderation analyses (how do intervention mechanisms interact with context to generate outcomes?).[19] Within systematic reviews, light might be shed on this by meta-regression or qualitative comparative analysis (what contextual factors are associated with increased effectiveness?).[20] Ideally, existing evidence will also indicate the size of effects, the cost-effectiveness of interventions and the balance between benefits and harms.

In the case of social and emotional learning for example, there is evidence from multiple studies that this intervention impacts on positive mental health and evidence from mediation analyses that this occurs via enhancement of social and emotional skills.[21 22] Furthermore, there is some evidence from

meta-regressions across studies that this is most likely to occur in US schools compared to those in other countries, and with older compared to younger elementary-age students.[23 24] There is also evidence that such interventions are cost-effective,[25] do not hamper academic education and in fact promote educational attainment.[26]

Secondly, we can be more confident that an intervention will generate effects in a new setting if there is local information which suggests that similar intervention mechanisms will be triggered which will generate similar outcomes. As suggested above, this is more likely when the current standard of treatment in the new setting is similar to that found in previous evaluation. It is also more likely when similar 'aetiological mechanisms' underlie risks in the new setting as those present in the settings of previous evaluation. Ideally, this would be assessed by comparing epidemiological evidence on the pattern of risk factors for the outcomes of interest in each setting.[4] In the case of social and emotional learning, we would compare the original and new setting in terms of a) the prevalence of deficits in social and emotional skills and b) whether these deficits appear to be risk factors for poorer mental health.[27]

In practice, comparable evidence may not always be available. Instead, the presence of similar aetiological mechanisms across the original and new setting may need to be inferred based on socio-demographic similarities between the populations in question. While this might generally provide a reasonable guide, such assumptions will sometimes be wrong, particularly where patterns of risk are in flux perhaps because of transformations in risk factors associated with major social or technological change. Hence, we suggest that, where there are concerns that aetiological mechanisms might differ, some local needs assessment is warranted to assess whether this might be so. Assessment might start with rapid qualitative assessments and move on to new epidemiological assessments of risk factors, depending on the degrees of uncertainty and urgency.[28]

Now let us consider scenarios where local implementation of a previously evaluation outcome is less likely to be feasible or generate similar outcomes.

Higher uncertainty scenario

Implementation

The first scenario is where there is uncertainty about implementation. Existing evaluations may not have been done in similar settings. They might have been done but identify barriers to implementation that may arise in the new setting, or identify facilitators that are unlikely to be present in the new setting. Evidence from local consultations or surveys, audits, guidelines for practice or observations might suggest important barriers to delivery. Considering the example of the school-based social and emotional literacy intervention: consultations or surveys with head- or classroom-teachers may suggest variable commitment to the intervention; audits of teacher skills might indicate deficits; analyses of school capacity might suggest that some teachers lack the time, autonomy, resources and/or senior support to deliver the intervention; or observations might indicate that classrooms are too disrupted by student misbehaviour to enable good implementation.[29]

In such scenarios, piloting of the feasibility of delivery before any full implementation would be indicated. Piloting might involve steps to address potential barriers to optimise implementation. It may well be possible to address challenges arising from provider capacity or other influences on implementation feasibility to ensure the intervention can be delivered locally. But piloting will need to assess this, monitoring fidelity and using interviews or focus groups to explore the local factors affecting feasibility. Local adaptations might be required.[5 6]

Effectiveness

Uncertainty about whether the intervention will be effective might occur for two reasons. Firstly, it might be due to uncertainty that the intervention will trigger the expected mechanisms. Existing evaluations may report effectiveness but be from settings or populations that do not closely resemble those found locally. Local information might suggest that the intervention will not be successful in triggering mechanisms locally. Consider the example of the school-based social and emotional education. Consultations with teachers might suggest that classroom-based interactions will be insufficient to develop students' social and emotional skills because there are concerns about student attendance or engagement. Consultation with parents or students might suggest that the social and emotional skills that the intervention promotes do not align with local social norms, again preventing lessons from improving these skills.

Secondly, there may be evidence that intervention mechanisms, even if triggered, will not generate the intended outcomes. This will be the case if different aetiological mechanisms are operating locally. Surveys or routine data might suggest: that rates of mental health problems among students are already low so there is little room for improvement; that deficits in social and emotional skills are uncommon; or that deficits in social and emotional skills are not significant risk factors for mental health problems. Needs assessments might suggest that, even if students develop social and emotional skills, other local aetiological mechanisms might swamp these intervention mechanisms: for example, if there are high rates of local violence or abuse that severely damage students' mental health. In such situations, intervention mechanisms may not disrupt the aetiological mechanisms underlying poor mental health. As above, some of the above forms of information may not already be available necessitating new research, starting with consultations and moving on to quantitative surveys if significant uncertainties remain and if the time and resources are available to enable this.

In terms of what actions are suggested, where the concern is about whether intervention activities will trigger mechanisms, this suggests the need at least for a new effectiveness trial potentially focused on an intervention adapted for local context. Where the concern is that the potential intervention mechanisms, even if triggered, will not disrupt local aetiological mechanisms or be swamped by other aetiological mechanisms, then local decision-makers may decide to seek entirely different interventions.

Discussion

We have aimed to develop guidance to help local decision-makers assess whether to implement interventions evaluated as effective elsewhere. We have stressed the importance of assessing existing evaluation evidence not merely in terms of what works overall but also: a) how successfully and with what reach interventions are implemented and what factors enable this (drawing on process evaluations); and b) what works for whom and how (drawing on mediation, moderation and/or qualitative comparative analyses within primary studies or systematic reviews).

We have contributed new insights to what is required of local needs assessments. First, we suggest that these should examine factors relating to the local potential, capability, contribution and capacity for implementation of the intervention. We recognise that such evidence will often not already exist and recommend the value of local consultations, surveys and audits in providing such information. Second, we have suggested that local needs assessment should where possible consider whether similar 'aetiological mechanisms' underlie adverse outcomes in a new setting compared with those of previous evaluations. We recognise that this can be challenging. Assessments about comparability should ideally be based on research assessing the prevalence and strength of risk factors for the outcomes of interest comparing the new setting to the settings of original evaluation. However, at minimum, similarities or differences in aetiological mechanisms

might instead be imperfectly inferred through similarities and differences in socio-demographic characteristics between populations.

We have suggested that where there is substantial uncertainty as to whether an intervention can be implemented in a new setting, this indicates the need for piloting the feasibility, acceptability and reach of implementation. Where there is uncertainty that implementation will trigger intended intervention mechanisms then this suggests the need for a new effectiveness trial. And when there is uncertainty about whether intervention mechanism, even when triggered, will generate the intended outcomes, this suggests that decision-makers may need to look to other types of intervention instead for their setting.

One limitation of our approach is that it is focused on scenarios where there is existing evidence on the intervention in question. In reality, there will be cases where an intervention is being considered for implementation in a setting when there is not evidence from multiple recent studies. In such cases, there will be much stronger arguments that new controlled trials of outcomes are needed before any delivery at scale should be considered. Furthermore, while we have recognised that evaluations must consider the size of effects, cost-effectiveness and the balance between benefits and harms, our various scenarios generally focus on simple questions of whether interventions are effective or not, in order not to over-complicate our analysis. While our focus has been on complex interventions and social/psychological mechanisms, we recognise that biological systems are complex and so our framework might also apply to clinical interventions, in an era of personalised medicine, co-morbidities and global pandemic.[30]

Of course, it might sometimes be decided to move to implementation of an intervention in a new setting even without firm evidence that this is locally appropriate, for example in response to emergencies arising from humanitarian crises and epidemics. In such cases, evidence might still be

collected in parallel to preparations for delivery. These might then inform refinements to the intervention and decisions about the extent and focus of accompanying evaluation and monitoring.[31]

References

1. Fletcher A, Jamal F, Moore G, et al. Realist complex intervention science: applying realist principles across all phases of the Medical Research Council framework for developing and evaluating complex interventions. *Evaluation* 2016;22(3):286-303.
2. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *Am J Public Health* 1999;89:1322-27.
3. Pawson R, Tilley N. Realistic evaluation: Sage 1997.
4. Bonell C, Oakley A, Hargreaves J, et al. Trials of health interventions and empirical assessment of generalizability: suggested framework and systematic review. *British Medical Journal* 2006;333:346-49.
5. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology* 2008;41:327-50.
6. Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *British Med Journal* 2004;328:1561-63.
7. Evans RE, Craig P, Hoddinott P, et al. When and how do 'effective' interventions need to be adapted and/or re-evaluated in new contexts? The need for guidance. *Journal of Epidemiology and Community Health* 2019;73(6):481-82.
8. Gardner F, Montgomery P, Knerr W. Transporting evidence-based parenting programs for child problem behavior (age 3–10) between countries: systematic review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology* 2016;45(6):749-62.
9. Loren T, Oliver K. Adverse effects of public health interventions: A conceptual framework. *Journal of Epidemiology and Community Health* 2014;68(3):288-90.
10. Kirke DM. Chain reactions in adolescents' cigarette, alcohol and drug use: similarity through peer influence or the patterning of ties in peer networks? *Social Networks* 2004;26:3-28.
11. Philiber S, Kaye JW, Herrling S. The National Evaluation of the Children's Aid Society Carrera Model Program to Prevent Teen Pregnancy. New York: Philiber Research Associations 2001.
12. Wiggins M, Bonell C, Sawtell M, et al. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ* 2009;339(b2534)
13. Freedman B. Equipoise and the ethics of clinical research. *New England Journal of Medicine* 1987;317:141–45.
14. Bhaskar R. The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences. Brighton: Harvester 1979.
15. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health*;38:65-76.
16. Evans R, Murphy S, Scourfield J. Implementation of a school-based social and emotional learning intervention: understanding diffusion processes within complex systems. *Prevention Science* 2015;16(5)):754–64.
17. Romasz TE, Kantor JH, Elias MJ. Implementation and evaluation of urban school-wide social–emotional learning programs. *Evaluation and Program Planning* 2004;27(1):89-103.
18. May C. Towards a general theory of implementation. *Implementation Science* 2013;8:18.
19. Bonell C, Fletcher A, Morton M, et al. 'Realist Randomised Controlled Trials': a new approach to evaluating complex public health interventions. *Social Science and Medicine* 2012;75(12):2299-306.

20. Thomas J, O'Mara-Eves A, Brunton G. Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. *Systematic Reviews* 2014;3:67.
21. Lewis KM, Bavarian N, Snyder FJ, et al. Direct and mediated effects of a social-emotional and character development program on adolescent substance use. *International Journa of Emotional Education* 2012;4(1):1-14.
22. Durlak JA, Weissberg RP, Dymnicki AB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development* 2011;82(1):405-32.
23. Goldberg JM, Sklad M, Elfrink TR, et al. Effectiveness of interventions adopting a whole school approach to enhancing social and emotional development: a meta-analysis. *European Journal of Psychology of Education* 2019;34:755–82.
24. Blewitt C, Fuller-Tyszkiewicz M, Nolan A, et al. Social and emotional learning associated with universal curriculum-based interventions in early childhood education and care centers: a systematic review and meta-analysis. *JAMA Network Open* 2018;1(8):e185727.
25. Belfield C, Bowden B, Klapp A, et al. The Economic Value of Social and Emotional Learning2015.
26. Corcoran RP, Cheung ACK, Kim E, et al. Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research. *Educational Research Review* 2018;25:56-72.
27. Thomson KC, Richardson CG, Gadermann AM, et al. Association of childhood social-emotional functioning profiles at school entry with early-onset mental health conditions. *JAMA Network Open* 2019 2(1):e186694.
28. Korteweg HA, van Bokhoven I, Yzermans C, et al. Rapid Health and Needs assessments after disasters: a systematic review. *BMC Public Health* 2010;10:295.
29. Herlitz L, MacIntyre H, Osborn T, et al. The sustainability of public health interventions in schools: a systematic review. *Implementation Science* 2020;15(1):4.
30. Matuchansky C. The promise of personalised medicine. *Lancet* 2015;386(9995):742.
31. Warsame A, Blanchet K, Checchi F. Towards systematic evaluation of epidemic responses during humanitarian crises: a scoping review of existing public health evaluation frameworks. *BMJ Global Health* 2020;5:e002109.

Figure 1: How implementation triggers intervention mechanisms which disrupt aetiological mechanisms to generate outcomes and how this is influenced by context