May 18th, 9:00 AM - May 21st, 5:00 PM

Commentary on Tone Kvernbekk's "Comparing two models of evidence"

David Hitchcock
McMaster University

Commentary on Tone Kvernbekk's "Comparing Two Models of Evidence"

DAVID HITCHCOCK
Department of Philosophy
McMaster University
Hamilton, Ontario
Canada
hitchckd@mcmaster.ca

1. Introduction

What is evidence-based practice? What is the structure of argumentation by which one defends a claim that one's practice is evidence-based? How does the reasoning go when one bases a decision about what to do on the evidence?

Answers to these questions are not obvious. Given the widespread acceptance in various professions and in policy-making that one's practices and policies should be evidence-based, the questions are worth exploring. That is what Tone Kvernbekk does in her paper, with specific reference to basing educational practice on the results of randomized controlled trials (RCTs). Her example is the decision of a teacher to give reading-delayed first-grade pupils extra phonological training, on the basis of an RCT in which children who received phonological training made significantly more progress than a control group on letter knowledge, single word reading and phoneme awareness.

Kvernbekk considers three models of the teacher's reasoning.

The first model is one in which the evidence is directly relevant to the teacher's decision. Kvernbekk takes this model to be implicit in Philip Davies' account (1999) of what constitutes evidence-based education, where the fifth and last step after organizing the evidence and grading its power is to determine its relevance to the educational practitioner's needs and environments. Proponents and critics of evidence-based education, she reports, share the assumption that basing practice on evidence involves deriving one's practice directly from the evidence. Kvernbekk argues, and I agree, that this assumption is mistaken. The direct source of the teacher's decision is the belief that the intervention will probably have the desired result. The evidence from the RCT plays an indirect role, somehow supporting this belief. So Kvernbekk sets aside this first model, and considers two other models that give the evidence an indirect role.

Her second model is Toulmin's well-known model for the layout of arguments. The teacher claims that (C) certain pupils in the first-grade class should get extra phonological training. The datum provided to support this claim is that (D) these pupils are reading-delayed. Kvernbekk is unclear about the warrant that would license the step from D to C; formally, it should be something like the warrant that (W) reading-delayed pupils in the first grade should get extra phonological training. As she points out, this warrant may be qualified by a qualifier such as (Q) 'presumably' and there may be corresponding conditions of rebuttal or exception, such as the hedge (R) 'unless they have extremely low initial scores on word recognition and letter knowledge.' The evidence from the RCT comes when the warrant is challenged, in the form of the backing that (B) in an RCT children who received phonological training made significantly more progress than a control group on letter knowledge, single word reading and phoneme awareness.

Bondy, P., & Benacquista, L. (Eds.). *Argumentation, Objectivity, and Bias: Proceedings of the 11th International Conference of the Ontario Society for the Study of Argumentation (OSSA), 18-21 May 2016*. Windsor, ON: OSSA, pp. 1-14.

The third model is Nancy Cartwright's model of evidence in use. Cartwright takes the conclusion of evidence-based practical reasoning to be a particular causal judgment, that an intervention *x* will likely work in a particular situation H (here)–in our example, that extra phonological training will likely help the first-grade teacher's reading-delayed pupils to read better. The direct support for this conclusion is a trio of premisses:

1. *x* is an INUS (insufficient but necessary part of an unnecessary but sufficient condition)[1] that works only if a, b and c are present.
2. a, b and c are present in situation H.
3. *x* can play a causal role in situation H.

Each of these premisses requires further support. In particular, premiss 3 requires support by a sub-premiss:

3.1 *x* played a causal role in situation T (there).

And this sub-premiss in turn gets support from the "evidence":

3.1.1 In situation T an RCT found a significant difference between the group that received intervention *x* and the control group that did not.

Thus on Cartwright's model the evidence on which the decision is based functions as a sub-sub-premiss in a sub-sub-argument of the reasoning to the conclusion that supports one's practice.

Kvernbekk notes that the latter two models are superior to the first one in giving the evidence an indirect role. She sees advantages in each model that the other model lacks. The Toulmin model speaks directly to the position of practitioners who notice something in their practice (the datum D, which Kvernbekk construes as practice-based evidence) and wonder what to do about it (C?), and it has a spiral structure that practitioners can navigate easily. The Cartwright model makes explicit the need to question whether what worked there will work here. Can one merge the models so as to incorporate the desirable features of each?

> Can we get a model that speaks to practitioners, has an internal logical drive, accommodates practice-based evidence, understands the role of RCT evidence as indirect, accommodates causes as INUS conditions, allows the restricted stretch of RCT designs and provides room for contextual evidence that takes us from there to here?

Kvernbekk notes that in general we cannot merge models by simply adding them together. Each must be tweaked in some way, and the merger will create a new third model, which she calls an 'ensemble'. To explore whether the Toulmin and Cartwright models can be merged in a way that preserves the advantages of each, Kvernbekk considers successively a number of ways of merging software models: matching, complementarity, slicing, cascading. She finds problems with each, and with their combination. One can match the sub-sub-premiss in

[1] The concept of an INUS condition comes from (Mackie 1965). Cartwright extends it to multi-valued variables with the following formulation: "X is an INUS contributor to Y: X is an insufficient but nonredundant part of a complex of factors that are unnecessary but together sufficient to produce a contribution to Y." (Cartwright 2012, p. 979)

Cartwright's model with the backing in Toulmin's model, whose content is the same RCT result, but the match may not be enough for a merger, given the different places of the matched elements. One can treat the models as complementary and combine the advantageous features of each, but the result seems like a jumble and lacks a place for Cartwright's sub-sub-premiss. Slicing off the third premiss of Cartwright's model with its supports would make it somewhat parallel to Toulmin's model, but would impoverish her model and have an uncertain fit with his. Cascading seems to work only when different models make distinctive contributions to a larger picture, but the Toulmin and Cartwright models are each to be understood as comprehensive. Kvernbekk concludes that, "if the purpose of our model merging is a more holistic view of argumentation which accommodates RCT evidence in an indirect role, the merging operations I have discussed fail to attain it"–a conclusion that she rather timidly characterizes as her "intuition".

2. Cartwright's account

Failure to find a solution does not mean that there is no solution to be found. Kvernbekk's "intuition" is thus more of a challenge than a discovery. And we should be encouraged to explore further by the fact that there is already a solution (Reed and Rowe 2005) to the general problem of translation between a Toulmin model of an argument and a Beardsley-Thomas box-arrow diagram of the sort that Cartwright uses.[2]

Before exploring whether the two models can be combined in a way that preserves the attractive features of each, we need to spell out a little more how the reasoning would go on Cartwright's model, whose diagram is obviously incomplete. Cartwright has been a vigorous critic for years of the inadequate advice given to practitioners and policy-makers on how to apply in a new setting (*here*, as she calls it) evidence that a certain causal intervention was effective in other settings (*there*, as she calls them). I propose to review the account she gave in her 2010 presidential address to the Philosophy of Science Association of what is needed to get from evidence of what was effective *there* to a well-supported prediction of what will work *here*. From Cartwright's requirements, I will then extract what I take to be her proposal for a sequence of reasoning that will take a practitioner or policy-maker from a consideration of evidence of effectiveness of some intervention to a well-supported prediction that it or its analogue will work in a new situation. I will model this reasoning in a box-arrow diagram, which I shall then convert into a Toulmin diagram using the sort of translation system proposed by Reed and Rowe. With both diagrams in place, we can then see whether there is a way of combining them that has the complementary features of each that Kvernbekk sees as advantageous. Finally, I will make some remarks about the point of Kvernbekk's project and propose a rather different approach to helping practitioners base their practice on evidence.

[2] Actually Cartwright uses the boxes without the arrows, stacking the premisses of the main argument for the ultimate conclusion (premisses that she very confusingly calls "major premises" [Cartwright and Hardie 2012, p. 17]) in a row of boxes just underneath the box at the top with the ultimate conclusion, and the premisses offered in support of each of the main premisses in a row below that row, and so on, without using any device to show which claim in the row immediately above a given box the premiss in that box is being used to support. Thus her diagrams are ambiguous in places. Also, she draws roots under ultimate premisses that do not need further support–a very confusing convention, since the word 'root' is usually used in such diagrams for what is at the top of the tree and the items at the bottom of each path are called 'leaves'. See figure 1.2 in (Cartwright and Hardie 2012, p. 17). Cartwright seems to have no knowledge of diagramming conventions that have been adopted for representing argument structure.

Cartwright describes as "often long and tortuous" (Cartwright 2012, p. 976) the road from evidence that some intervention *x* works somewhere to the conclusion that it will work here. She argues that building this road needs four essential materials:

1. laws wide enough to cover both the evidence and the prediction.
2. supports without which *x* cannot act.
3. ladders for climbing up and down levels of abstraction.
4. unbroken bridges by which the influence of the cause can travel to the effect.

Let us consider each of these in turn.

2.1. Laws

An ideally designed RCT, which is approximated in real life by the studies given the highest grade in one of the clearinghouses that review such studies, will establish with high probability that the one factor *x* that was present in the intervention group but not the control group contributed causally, *in the circumstances of the trial*, to the differential occurrence or value of outcome *y* in its intervention group. Causal contributors of this sort are usually not sufficient causes by themselves, but work in combination with other factors in the situation to produce the desired outcome (or desired change in value) in some individuals in the study. In the simplest case, there is a deterministic linear relationship between the value of the outcome of interest in each individual *i* in the study and a combination *b* of causally relevant contributors to that outcome, a relationship that can be expressed in the equation:

$y(i)c = a_1 + a_2y_0(i) + a_3b(i)x(i) + a_4z(i)$ (Cartwright and Hardie 2012, p. 26),

where:

- $y(i)$ is the value of the outcome *y* for individual *i* in the study group,
- '$c =$' indicates that the value on the left is equal to the value on the right as a result of the causal contributions of the factors joined by the plus signs,
- the *a*s are constants across all individuals,
- a_1 is a constant in the situation summarizing the net effect of all those factors that have an influence on the value of the outcome independently of differences between the individuals *i*,
- $y_0(i)$ is the value of *y* for individual *i* at the start of the study,
- $b(i)$ sums up the values for each individual *i* in the study of the necessary factors that work along with the intervention factor *x* to make a combination that is sufficient to influence the value of *y* for individual *i*,
- $x(i)$ is the value of the intervention variable for individual *i*, and
- $z(i)$ sums up the values for individual *i* of all the INUS conditions affecting the value of *y* for individual *i* that do not include *x*.

An RCT, if carried out according to the ideal, holds all factors constant between the intervention and the control group except the value of factor *x*. If the average values of y in the two groups

differ by an amount that is statistically significant (i.e. probably not due to chance), then we can be almost sure that *x* is part of some complex that is causally sufficient to change the value of *y* in at least some individuals in the study, since the other factors in the equation cancel out when the averages are subtracted (pp. 976-977).[3] To be justified in applying the intervention *x* in a new situation, one needs the assumption that the law expressed in the equation just mentioned, or some modified version of it that includes *x*, applies in the new situation, and one should have good reason to make this assumption. Simple induction, whether by direct extrapolation or through an intermediate generalization, needs a stable principle whose form is that of the equation, as well as support for this principle by a stable substructure of factors like those mentioned in the equation.

Cartwright appears to say little about how the first two factors in her equation should be taken into account by a practitioner deciding to make a similar intervention in a new situation. Sometimes (e.g. Cartwright 2012, p. 976) she even omits the second factor altogether in her description of the form of the causal law that is at work. This second component is the place where the practitioner's identification of the educational situation *here* comes into play in her model, even though she does not explicitly mention it in her account of reasonable evidence-based policy-making.

2.2. Supports

The supports are the values of *b* in the equation that work together with the values of *x* so that the combination is a contributor to the value of *y*. As Cartwright points out, these values are not constant across the members of the study, and the contribution of the combination may differ from one individual to another, in some cases even making a negative contribution. A practitioner using the results of a study as the evidential basis for doing something similar will want to pick those values of the support factors that were most helpful in combination with factor *x* in producing the desired outcome. For this purpose, subject-matter knowledge is clearly required.

To underline the need for support factors, Cartwright cites a failure in evidence-based policy-making in education. A randomized control design had found that in Tennessee students in smaller classes did better at all K-3 grade levels than students in larger classes, and that minority and inner-city children gained two to three times as much from reduced class sizes as white and non-urban peers. The study confirmed what previous meta-analyses of less rigorously designed studies had shown and what common sense and popular opinion believed. When in the 1990s class sizes were reduced in California, however, there was no conclusive link between class size reduction and student achievement, and no greater effect among disadvantaged children (Cartwright and Hardie 2012, p. 4). The replication failed because, unlike Tennessee, California lacked the spare space and qualified teachers needed for the larger number of reduced-size classes. It hired in a short time a large number of teachers, many of whom were unqualified. And it took space that was being used for other programs that contributed to student achievement: special needs, music and arts, athletics, and child care programs (Cartwright and Hardie 2012, p. 65). Cartwright and Hardie draw two general morals from these causes of

[3] Cartwright (2011, p. 223) claims that the argument is deductively valid once one adds the assumption that outcome differences have a cause, but even statistically significant differences can be due to causally relevant differences between the intervention and control group that randomized allocation did not prevent from occurring. Such an explanation is just very unlikely.

failure. First, the practitioner needs to do a "horizontal search" to determine the support factors (the *b* factors in the equation) that are needed to work with the intervention to make it effective; in this case, a required support factor was qualified teachers. Second, the practitioner needs to consider what factors independent of the intervention (the *z* factors in the equation) make a causal contribution to the desired result and to make sure that implementing the intervention does not undermine those other factors with a resulting nullification of its effect; in this case, activities that had to be abandoned in order to open space for extra classes were making an independent contribution to student achievement, which the reduction in class sizes eliminated.

2.3. Ladders

What Cartwright calls 'ladders' are principles that enable one to move from one level of abstraction to another. A practitioner or policy-maker might want to describe the intervention variable more abstractly than it was described in the study establishing its effectiveness, so as to make it applicable to a new situation. To do so is to move up the level of abstraction by making an assumption about what it was about the intervention that contributed to the desired outcome. If one contemplates using a somewhat different concretization of this abstractly described intervention in the new situation, one moves back down the ladder of abstraction. The issue in this case is whether a different concrete realization actually constitutes the same type of abstractly described intervention that was supposed to be the operative factor in the study.

Cartwright illustrates the importance of abstracting the causal principle through a spectacular failure of evidence-based policy-making. In the Indian Tamil Nadu Integrated Nutrition Project (TINP), a combination of supplementary feeding for children under 24 months, health measures and education of pregnant mothers produced a significant decline in malnutrition in the rural areas of districts in Tamil Nadu state. The same intervention in Bangladesh had little success. The reason was that the food provided was used as a substitute rather than a supplement, with the food that would have been used by the mother and the young children given to others; in Bangladesh it is not the mother who does the shopping but the father, and in joint households where the mother lives with her mother-in-law it is the mother-in-law who controls the women's domain (Cartwright and Hardie, p. 82). The causal principle that was operative in Tamil Nadu, that "Better nutritional knowledge in mothers plus food supplied by the project for supplemental feeding improves the nutritional status of their children" (Cartwright 2012, p. 983), simply did not hold in Bangladesh, because of differences in who controlled the selection and distribution of food for a family. To get a principle that explains the positive results in Tamil Nadu but also applies in Bangladesh, we need to describe the components of the sufficient causal condition in Tamil Nadu more abstractly, as follows:

> Better nutritional knowledge results in better nutrition for a child in those who (a) provide the child with supplemental feeding, (b) control what food is procured, (c) control how food gets dispensed, and (d) hold the child's interests as central in performing b and c. (Cartwright 2012, p. 984)[4]

Implementing this causal principle in Bangladesh would require a different way of concretizing its abstractly described causal factors.

[4] The formulation is syntactically awkward. More grammatically: A child will get better nutrition as a result of improving nutritional knowledge in those who meet conditions a through d.

2.4. Unbroken bridges

Often the input variable *x* contributes in the study situation to a change in the outcome variable *y* through a chain of intermediate causes. For simplicity, let us consider a chain with just one intermediate variable *u*. Replication of the values of *x* in a new situation may not have the same outcome because the value of the intermediate variable *u* is a realization of a different abstract property in that situation, one that has a different abstract effect. In the study, the value of *x* makes a contribution in virtue of some factor **X**, which causes immediately some effect **U** of which the value of *u* is a concrete realization, and factor **U** causes some effect **Z** of which the value of *z* is the concrete realization. We have to go up the ladder of abstraction to discern the causal principles at work. In the new situation, however, the same value of *u* may constitute a different abstractly described causally operative feature **U′**. And **U′** may contribute to a completely different outcome **Z′**, with a quite different concrete realization than the intended value of *z*.

Cartwright illustrates the need for "unbroken bridges" with a hypothetical example of failure. The United Kingdom heavily encourages a child's caregivers to attend parenting classes. The justification for this policy is that such encouragement will increase attendance at such classes, which will make caregivers more knowledgeable about how to care for their child, which will improve the child's welfare. However, some social workers worry that in certain cultural groups fathers pressured to attend parenting classes will interpret their coerced attendance as public humiliation, which will make them more violent, perhaps to the child (Cartwright 2012, p. 987). Cartwright interprets the problem as one of variant abstract descriptions of attendance at classes, as compliance and as humiliation. The lesson she draws is that, if there is a chain of causation, one needs to take into account the abstract descriptions under which each link in the chain might have effects in a given situation, in order to make sure that the desired effect occurs.

3. Cartwright's model

In her co-authored book on evidence-based policy-making (Cartwright and Hardie 2012), Cartwright ignores the sub-argument in her model from the results of a study to the conclusion that an intervention *x* was effective in securing a desired outcome *y*. As Cartwright and Hardie (2012, p. 54) point out, practitioners can get trustworthy information on the so-called "internal validity" of studies of causal effectiveness from the clearinghouse that reviews and grades studies relevant to their field. Cartwright's focus is on the main argument from the premiss that "*x* worked *there*" to the prediction that "*x* will work *here*," in the practitioner's situation. She points out that one can gloss "*x* will work *here*" in various ways, some of which require more support than others, but that a minimally demanding gloss that is satisfactory for most practitioners' purposes is that *x* will play a positive causal role *here* for at least some individuals (Cartwright and Hardie 2012, pp. 41-42). The main argument for such a prediction, which she calls "the effectiveness argument," goes as follows:

1. *x* played a positive causal role there.
2. *x* can play the same role here after it is implemented.
3. The support factors necessary for *x* to play a positive causal role here after *x* is implemented are present for at least some individuals here.

> C. *x* will play a positive causal role here for at least some individuals. (Cartwright and Hardie 2012, p. 45)[5]

Understanding this argument scheme requires understanding the concepts of causal role and support factor. Cartwright defines first the concept of *being able* to play a causal role with respect to a specified outcome in a specified situation:

> *x can play a causal role with respect to outcome y* in a situation if *x* genuinely appears in the causal principles [with outcome *y*–DH] for that situation. (Cartwright and Hardie 2012, p. 44; italics in original)

She uses that definition in defining the concept of *actually* playing a causal role with respect to a specified outcome in a specified outcome:

> *x does* (or *will*) *play a causal role* [*with respect to outcome y*–DH] *in situation S* if it can play a causal role [with respect to outcome *y*–DH] under the [causal–DH] principles [with outcome *y*–DH] that govern S and the support factors (designated by *b*) required under those principles take nonzero values for some individuals. (Cartwright and Hardie 2012, p. 44; italics in original)

A support factor is thus an INUS condition that appears along with the intervention *x* in a causal principle with the outcome *y* that is the practitioner's concern.

On Cartwright's analysis, the crucial steps for a practitioner wanting to base practice on evidence that an intervention had a desired outcome somewhere else are the sub-arguments in support of the second and third premisses of the effectiveness argument, which we might label for short the causal role premiss and the support factors premiss. Both premisses require for their support some insight into *how* the intervention *x* contributed to the desired outcome in the situations where it has been shown to work. The system of grading studies of causal effectiveness for their "internal validity" tells us nothing about that question. One needs subject-matter knowledge rather than a formal methodology to answer it, and this subject-matter knowledge needs to be applied to the descriptions of the studies that have been graded as establishing with reasonable confidence a causal relationship in the study subjects between intervention *x* and outcome *y*.

The sub-argument for the causal role premiss (that *x* can play the same role *here* after it is implemented as it played *there*) requires what Cartwright and Hardie (p. 91) call "vertical search": describing the causal principle by which *x* contributed *there* at a high enough level of abstraction that it can reach the situation *here*, and then going down the ladder of abstraction to figure out what that principle amounts to *here*. At the same time the principle must not be described so abstractly that it becomes difficult to move back down the ladder of abstraction to the particularities of the situation *here*. As they point out, any good evidence of the effectiveness of some intervention *x* in contributing to some positive outcome *y* can be used as the basis of the recommendation: "To solve any problem, give the right resources to the right agents in the right

[5] I have changed the wording slightly so as to make the meaning clearer. A variant with "the policy" in place of "*x*" appears on page 54.

circumstances" (Cartwright and Hardie 2012, p. 86). But this recommendation is far too abstract to generate any specific advice in a specific situation.

The sub-argument for the support-factors premiss (that at least some individuals *here* have the support factors needed for *x* to play a positive causal role here) involves what Cartwright and Hardie (2012, p. 91) call "horizontal search": determining whether the support factors in the operative causal principle as it has been concretized for the present situation are actually present, or at least could be created.

Cartwright's main message is that the practitioner or policy-maker needs not just evidence *that* a particular intervention worked elsewhere but an understanding of *how* it worked. This understanding requires figuring out how the intervention combined with other factors in the situations where it has been shown to "work" to contribute causally to the desired outcome, as well as what other combinations of factors not involving the intervention contributed positively or negatively to the desired outcome. The causal principles so identified then need to be reframed in a more abstract way to the extent necessary to make them applicable to the practitioner's or policy-maker's situation. These abstractly formulated principles then need to be concretized for that situation, so as to identify first what the analogous intervention would amount to in the new situation, then whether the other components of the concretized principle that includes the intervention are present or can be made to be present in the new situation, and finally whether the other principles as concretized for the new situation would continue to be operative if the analogous intervention were implemented. Only positive answers to the latter two questions would justify implementing the analogous intervention in the new situation.

4. Fleshing out Cartwright's model

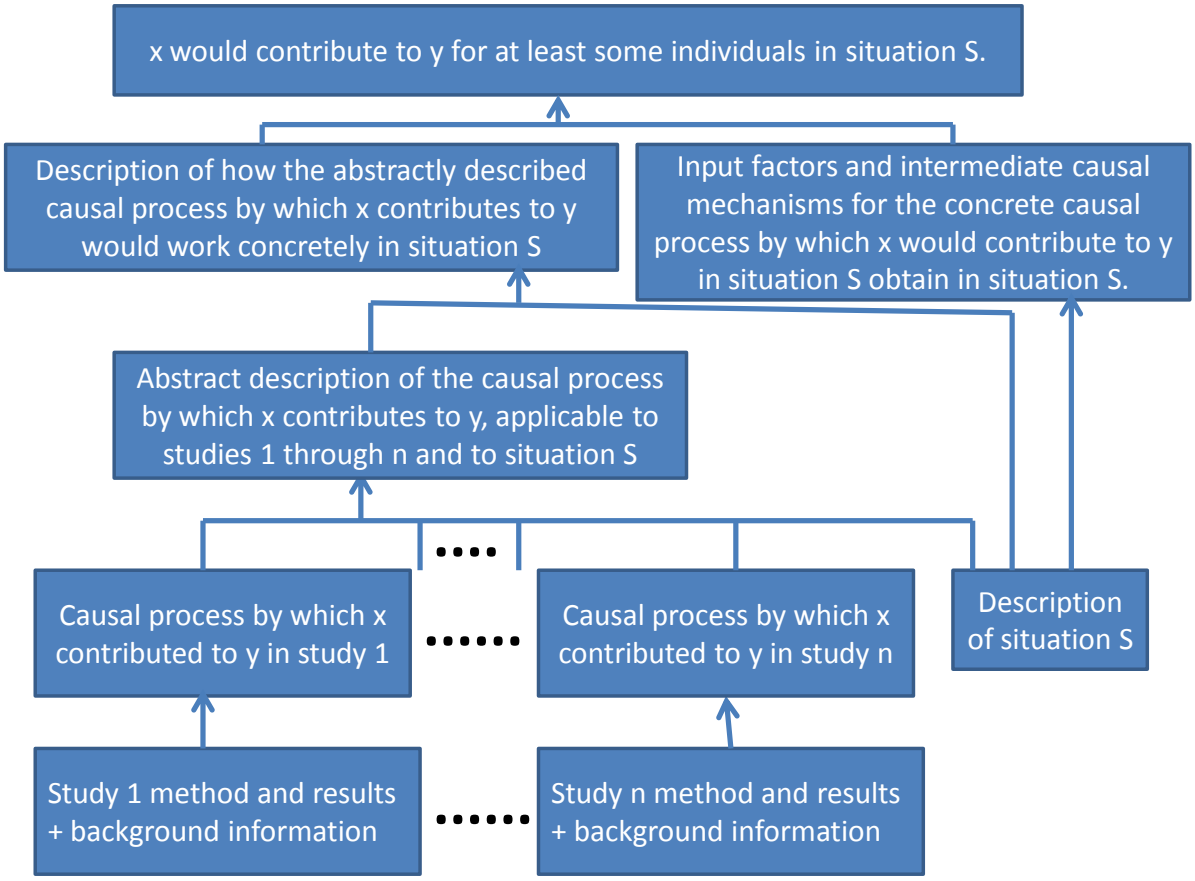
Cartwright never gives her readers a fully worked out example or fully worked out scheme for this chain of reasoning. Her pyramid diagrams include boxes for the main conclusion, main premisses, sub-argument premisses and sub-sub-argument premisses, with roots at the bottom of premisses taken not to need an argument in support of them (Cartwright and Hardie 2012, p. 17). But her diagrams modelling the reasoning that an intervention will work *here* have many boxes whose content is designated only with a question mark (Cartwright 2011, p. 223; Cartwright and Hardie 2012, p. 133). Further, when one thinks through how the reasoning that she recommends would actually go when one starts with evidence that an intervention worked somewhere and ends with the conclusion that it will work in a specified situation, one discovers that the content that she does provide in her incomplete pyramids does not correspond to the content that one gets by putting her recommendations into a box-arrow diagram.

We can start to construct a box-arrow schema for reasoning from evidence to practice by spelling out schematically in words the steps that Cartwright recommends:

1. For each study or practice-based experience that provided good evidence that a particular type of intervention contributed to the desired outcome, combine the description of its method and results with domain-specific background information to get as complete a picture as feasible of the causal process by which the intervention contributed to the outcome.
2. With reference to the situation in which one is contemplating using a similar intervention to contribute to the same kind of desired outcome, formulate the descriptions of the causal process that emerged from step 1 at the lowest level of abstraction that covers both the studies and experiences considered at step 1 and the new situation.

3. Work out what the abstractly described intervention and other inputs to the causal process in which it contributes to the desired outcome amount to concretely in the new situation.
4. Determine whether the other inputs exist or can be made to exist in the new situation.
5. Determine whether the causal process is likely to unfold in the way described at step 3 if the analogous intervention is implemented and the other inputs to the causal process exist.
6. If the answers at steps 4 and 5 are affirmative, conclude that the analogous intervention will contribute to the desired outcome if it is implemented in the new situation and any other inputs to the causal process that are not already present are provided.

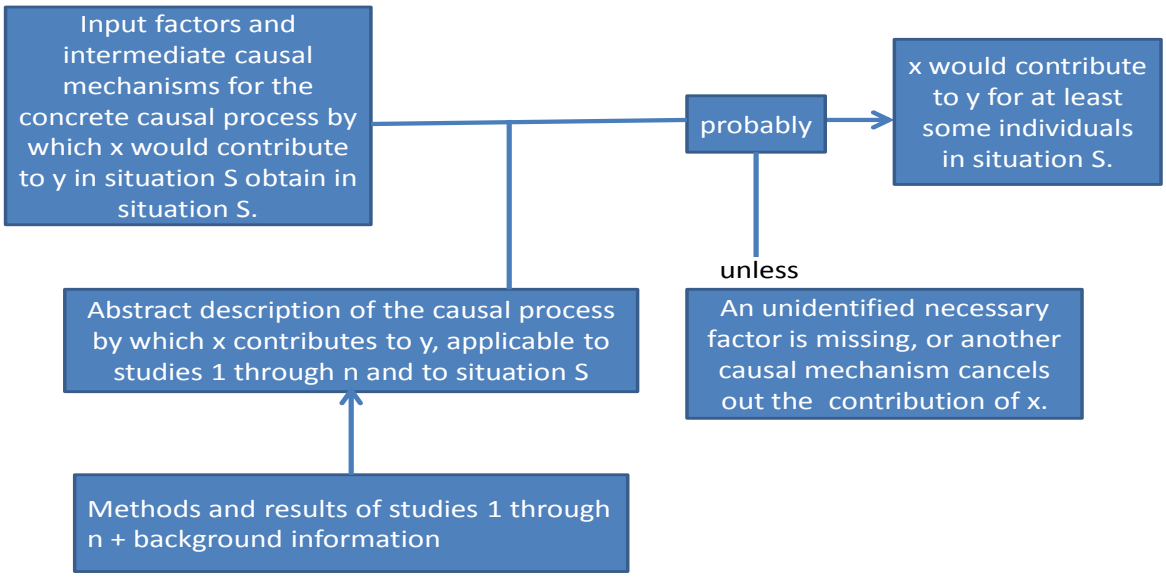
One can create a box-arrow diagram for the chain of reasoning involved in these steps, as follows:



It should be noted that further reasoning is required to get to a decision to implement the analogous intervention in the new situation. The reasoning involved is an example of means-end reasoning, which in general requires in addition to the premiss that a contemplated means will be effective in contributing to a desired goal premisses that the means is permissible, no less effective than any alternative means, and free of counter-balancing negative side-effects (Hitchcock 2011).

The box-arrow diagram can easily be converted into a Toulmin-style diagram, with some telescoping. The descriptions on the bottom row of the method and results of the studies showing that *x* contributed to *y*, along with the background information, constitute the backing for the abstract description, applicable to studies 1 through n and to situation S, of the causal process by

which *x* contributes to *y* in all these situations. That abstract description is the warrant for an inference from the data about the input factors and intermediate causal mechanisms to the conclusion that *x* would contribute to *y* for at least some individuals in situation S. The qualifier and rebuttal in the Toulmin model enable one to hedge one's conclusion and to mention factors that might block the inference to the conclusion.



5. Comparing the two models

We have now found a way to translate a box-arrow diagram of reasoning from evidence of effectiveness *there* to a prediction of efficacy *here* into a Toulmin diagram of the same reasoning, and vice versa. Each diagram includes information that the other lacks. The box-arrow diagram has more detail about the process of identifying the specific causal process in each study that showed that an intervention *x* was effective in contributing to a desired outcome *y*, then moving to an abstract description of that process that is applicable to the study situations and to the practitioner's situation, then concretizing that abstract description for the new situation. The Toulmin diagram flags the qualified support for the prediction and draws attention to conditions of exception that might obtain in the new situation. However, each diagram could be expanded so as to include the information that it currently lacks.

We can now return to Kvernbekk's question:

> Can we get a model that speaks to practitioners, has an internal logical drive, accommodates practice-based evidence, understands the role of RCT evidence as indirect, accommodates causes as INUS conditions, allows the restricted stretch of RCT designs and provides room for contextual evidence that takes us from there to here?

To my mind, the Toulmin diagram would be more readily grasped by practitioners than the box-arrow diagram. The practitioner starts with a desire to affect in a positive way some variable *y*, such as the reading skills of beginning readers. On the model of evidence-based practice, this desire should lead to consultation of a clearinghouse that rates studies of the effect on *y* of various interventions, where one or more highly-rated studies will show that some intervention *x* had the desired positive effect. To get the backing in the Toulmin diagram, the practitioner needs to combine the description of method and results in these studies with practice-based and theory-based background information to get a description of the causal process by which *x* has been found to positively affect *y*, one that is abstract enough to cover both the study situations and the practitioner's present situation. With this warrant available, the practitioner can then check for the concrete realization in the present situation of the factors involved in the causal process. Having found them, the practitioner can then make a qualified prediction that an intervention analogous to *x* will have a similar effect in the present situation. The qualifier reminds the practitioner that the prediction is not established with certainty, and the rebuttal alerts the practitioner to check for possible missing unidentified necessary factors or possible overriding causal processes created by the intervention. Thus there is a logical drive–from the desire to affect *y* to the backing to the warrant to the data to the conclusion and then to the qualifier and the rebuttal. Practice-based evidence is accommodated in the backing and in the data. The RCT evidence is in the backing, and thus has an indirect role. The role of *y* as an INUS condition can be accommodated in the warrant; however, its role may be more complex, if there is a chain of causality. A diagramming of causal processes like that advocated by Judea Pearl (2009) might be a more useful tool than Cartwright's cakes (as at Cartwright and Hardie 2012, pp. 62 and 65), which make no provision for chains. The restricted stretch of RCT designs is not explicitly flagged in the Toulmin model, but could be marked by including some of the intermediate steps in the box-arrow diagramming of reasoning from evidence of effectiveness to a prediction of efficacy. The contextual evidence that takes the practitioner from the situations where an intervention was found to be effective to a new situation is accommodated in the data used as the basis of the prediction.

The box-arrow diagram has most of the features that Kvernbekk seeks in a model of reasoning for evidence-based practice. But it lacks the key feature of having a logical drive. Looking at the hierarchical structure, a practitioner would have little sense of where to start. So, as far as I can see, for this purpose the Toulmin model wins.

6. Another approach

One may wonder whether such a model is the best way for practitioners to start thinking about how to improve outcome *y* in the particular situations in which they find themselves. An argument diagram represents an argument that has already been constructed, and is a tool for checking how well the argument supports its ultimate conclusion. The process of deciding on an intervention with a view to securing a desired outcome is a process of discovery, not a process of justification. The argument that supports the decision at which one arrives is something constructed after one has arrived at it. For help in the process of discovery, one needs something like a problem-solving or critical thinking checklist of steps to be gone through. It should be a checklist rather than a sequence, not only to allow for variation in the order in which one checks off the various steps but also to allow for recycling among the steps. For example, one step in the process is to define the problem, e.g. as that of helping reading-delayed children at the end of

their first year of reading instruction to improve their reading skills. As one goes through other steps in the process, such as that of gathering research-based and practice-based evidence of what can work to solve one's problem, one may be led to redefine the problem. The redefinition of the problem may in turn motivate a new search for evidence on what can work.

We can take Cartwright's set of requirements for a well-supported evidence-based prediction as a guide to the questions that a practitioner should have in mind when deciding on an intervention aimed at securing a desired outcome. The following is a preliminary sketch:

- What is the problem? (What desired outcome, not yet achieved, do I wish to bring about?)
- What interventions have been found, either in research or in practice, to solve this problem?
- How effective was each of these interventions found to be?
- How good is the evidence that each of these interventions worked?
- What is the causal mechanism by which each of these interventions worked?
- How can one describe this causal mechanism abstractly enough that it applies to my situation?
- What concretely would the ingredients of the abstractly described causal mechanism be in my situation?
- Is it feasible to implement these ingredients in my situation?
- What are the likely side effects of implementing these ingredients in my situation?
- On balance, which one of the feasible interventions is likely to be effective and without adverse side effects?

The use of bullet points rather than numbers is meant to avoid the impression of an invariant sequence. For example, it may be obvious right away that an intervention that has been found to work is just not feasible in the practitioner's situation, so that one can simply discard that intervention from further consideration. Or one may realize right away that a feasible and effective intervention, such as making reading-delayed students repeat the first year of reading instruction, has side-effects so adverse that it should be dropped from consideration.

The need to understand the causal mechanisms at work in an intervention found to be effective suggests an agenda for improvement of the information available in clearinghouses that review and grade studies of the effectiveness of interventions. If they do not already do so, they need to supplement their grading with a description of the entire causal mechanism by which the intervention does its work, formulated at a high enough level of abstraction that practitioners and policy-makers can concretize it for their own particular situation. These descriptions should be subject to peer review and debate, with acknowledgement of uncertainties and of alternative explanatory models. The isolated practitioner or policy-maker cannot reasonably be expected to go through the complicated work of reconstructing a causal process from the mere grading of studies for their methodological rigour in ruling out alternative explanations of a statistically significant difference between an intervention group and a control group. Randomized controlled trials in particular treat the other components of the process by which an intervention secures its effect as a black box. But the decision to conduct a randomized controlled trial generally rests on preliminary evidence that an intervention is likely to work, evidence that includes an understanding of the causal process by which it would have its effect. And generally both research-based and practice-based evidence of what can work includes some understanding of

how it has been effective. In the case of research-based evidence, clearinghouses have the information and the resources to articulate this understanding in a way that helps practitioners and policy-makers. They should do so. It is not enough to recognize, as is done for example in the system of grading studies in evidence-based medicine, that research evidence of an intervention's effectiveness can sometimes have only indirect relevance to a practitioner's decision-making (Guyatt et al. 2011). Especially in the case of fields like education and development economics, where social, psychological and cultural factors often influence the effectiveness of interventions, practitioners and policy-makers need an understanding of the causal principles at work.

References

Cartwright, N. (2011). Predicting what will happen when we act. What counts for warrant? *Preventive medicine* 53 (4), 221-224.

Cartwright, N. (2012). Will this policy work for you? Predicting effectiveness better: How philosophy helps. *Philosophy of Science* 79 (5), 973-989.

Cartwright, N., and J. Hardie (2012). *Evidence-Based Policy. A Practical Guide to Doing It Better*. Oxford: Oxford University Press.

Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies* 47 (2), 108-121.

Guyatt, G. H., A. D. Oxman, R. Kunz, J. Woodcock, J. Brozek, M. Helfand, P. Alonso-Cello, Y. Falck-Ytter, R. Jaeschke, G. Vist, E. A. Akl, P. N. Post, S. Norris, J. Meerpohl, V. K. Shukla, M. Nasser and H. J. Schünemann (2011). GRADE guidelines: 8. Rating the quality of evidence—indirectness. *Journal of Clinical Epidemiology* 64 (12), 1303-1310.

Hitchcock, D. (2011). Instrumental rationality. In: P. McBurney, I. Rahwan and S. Parsons (Eds.), *Argumentation in Multi-Agent Systems* (pp. 1-11). Berlin: Springer.

Kvernbekk, T. (2016). Comparing two models of evidence. Paper presented at a conference on argumentation, bias and objectivity sponsored by the Ontario Society for the Study of Argumentation at the University of Windsor, May 18-21, 2016.

Mackie, J. L. (1965). Causes and conditions. *American Philosophical Quarterly* 2, 245–264.

Pearl, J. (2009). *Causality*. New York: Cambridge University Press.

Reed, C., and G. Rowe (2005). Translating Toulmin diagrams: Theory neutrality in argument representation. *Argumentation* 19 (3), 267-286.