

This item is the archived peer-reviewed author-version of:

The ResQ approach : theory building across disciplines using realist evaluation science and QCA

Reference:

Renmans Dimitri.- The ResQ approach : theory building across disciplines using realist evaluation science and QCA
International journal of social research methodology - ISSN 1464-5300 - 26:4(2023), p. 469-482
Full text (Publisher's DOI): <https://doi.org/10.1080/13645579.2022.2052695>
To cite this reference: <https://hdl.handle.net/10067/1869280151162165141>

1 **The ResQ approach: Theory building across disciplines using realist** 2 **evaluation science and QCA**

3 Published in International Journal of Social Research Methodology, <https://doi.org/10.1080/13645579.2022.2052695>
4 Do not use page numbers for referencing.

5 The last decades have seen an enormous growth in published research and
6 evaluations, which makes it difficult for a researcher to stay up-to-date in their
7 own field, let alone complement their knowledge with insights from other fields.
8 In this paper we give an elaborate overview of a methodology that aims to tackle
9 this task. It builds on the realist evaluation science approach and combines it with
10 qualitative comparative analysis (QCA), hence its name: the ResQ approach.
11 Central to the approach are generative mechanisms that can be found across
12 fields, domains, sectors and contexts. The approach sets out to synthesize the
13 evidence on the circumstances linked to the triggering of these mechanisms.
14 QCA is used to identify the most relevant conditions, leading to theories around
15 these mechanisms, called ‘mechanism concepts’. New studies can test, and refine
16 the mechanism concepts, setting up a continuous cycle of theory-building across
17 disciplines enabling us to learn from other fields, disciplines and contexts in a
18 systematic way.

19 **Keywords:** realism; mechanism; research synthesis; theory building; qualitative
20 comparative analysis

21 22 **Introduction**

23 According to the latest annual report by Crossref, as of October 2019, almost 80 million
24 journal articles have a DOI, which is an increase of 8% compared to the preceding year
25 of 2018 (Crossref, 2019). This means that every year academics are being flooded with
26 hundreds of papers in their own fields alone. Staying up-to-date is a daunting task, let
27 alone finding the time to complement knowledge from one’s own field with knowledge
28 from other fields. In this paper a methodology is presented capable of bringing all this
29 information together making this task more manageable, stimulating cross-discipline
30 learning, and advancing science and policy-making.

31 The proposed methodological approach is based on the ‘realist evaluation
32 science’ methodology put forward by Pawson (2013), as it is ‘sympathetic to the usage
33 of a multi-method, multidisciplinary evidence base’ (Pawson et al., 2004, p. iii).
34 Moreover, as Emmel (2021) rightly argues ‘a pre-requisite of a realist methodology
35 must be that it moves beyond the limits of institutionalised disciplinary science’ (p. 95).
36 Finally, at the centre of the realist approach is theory building (Pawson, 2013), which
37 has been noted to be very useful for combining knowledge from different disciplines
38 (Abner et al., 2017; Perry, 2010).

39 Realist evaluation aims to be a bridge between evaluation and science. It is
40 inspired by Bhaskar’s (2008 [1975]) critical realism, hence has a generative view of
41 causality. This means that phenomena and outcomes of interventions are caused by
42 underlying and most often unobservable mechanisms (Astbury & Leeuw, 2010;
43 Bhaskar, 2008 [1975]; Pawson, 2013). Realist programme theories describe the specific
44 contextual conditions linked to the triggered mechanisms and the interactions between
45 mechanisms that produce specific outcomes.

46 To strengthen the analysis of these conditions, the in this paper presented ResQ
47 approach combines an adaptation of the realist evaluation science approach with
48 qualitative comparative analysis (QCA) (Ragin, 2014 [1987]). The latter is an approach
49 and technique used to compare cases and discern the necessary and sufficient conditions
50 for an outcome to occur.

51 In this way the ResQ approach aims to synthesize research and evaluation
52 findings and build theories in a similar way as Pawson’s (2013) ‘reusable conceptual
53 platforms’. These theories can inform policy makers by explaining how interventions
54 work and how to make them more effective.

55 This paper presents the underlying rationale of the ResQ approach and its
56 constitutive elements, elaborates on the different steps and illustrates this with a short
57 pilot example in the appendix. We now turn to a discussion of the two approaches on
58 which the ResQ approach is based before elaborating on the latter.

59 **The realist approach**

60 The realist approach developed by Pawson and Tilley (1997) is a theory-driven
61 approach to evaluation and research synthesis. This means that every evaluation or
62 synthesis begins with an initial programme theory which guides the analysis, and ends
63 with a refined programme theory. Rather than responding to the question of whether an
64 intervention was effective, the realist approach attempts to answer the more nuanced
65 question: ‘What works for whom, when, where and why?’ (Pawson, 2013).

66 What makes realist evaluation different from other theory-driven approaches is
67 the philosophical influence of scientific and critical realism (Bhaskar, 2008 [1975];
68 Pawson, 2013; Sayer, 2000). Consequently, it adheres to generative causation, meaning
69 that outcomes are generated or caused by mechanisms that can be defined as
70 ‘underlying entities, processes, or structures’ (Astbury & Leeuw, 2010, p. 368). Indeed,
71 according to realism, reality is ontologically stratified in the domains of the real, the
72 actual and the empirical (Bhaskar, 2008 [1975]). The domain of the real is comprised of
73 mechanisms and structures and exists independent of our knowledge of it. These
74 mechanisms and structures generate events (domain of the actual) which we can
75 observe through our experiences (domain of the empirical) (Bhaskar, 2008 [1975];
76 Byrne, 2018). When trying to explain an observed relationship within a system (i.e.
77 between an intervention or a context and an outcome) - which occurs in the domain of
78 the actual and is measured or observed in the domain of the empirical - mechanisms and

79 structures from the domain of the real ‘tell us what it is about that system that generates
80 the uniformity’ (Pawson, 2006b, p. 23).

81 Thus, ‘[m]echanisms explain causal relations by describing the “powers”
82 inherent in a system’ (Pawson, 2006b, p. 23). Evidently, this also means that these
83 mechanisms are linked to ‘particular contexts to generate outcomes of interest’ (Astbury
84 & Leeuw, 2010, p. 368). To emphasize this link with context, realists use a heuristic
85 called the ‘context-mechanism-outcome configuration’ or CMOC. This means that a
86 specific context leads to a specific outcome due to a specific mechanism. These
87 CMOCs take centre stage in the realist approach and are used as a heuristic to structure
88 the programme theory and the analysis.

89 As mentioned, a realist approach begins by developing an initial programme
90 theory, which is at what Merton (1968) called the middle-range level. Middle-range
91 theories are ‘theories that lie between the minor but necessary working hypotheses that
92 evolve in abundance during day-to-day research and the all-inclusive systematic efforts
93 to develop a unified theory that will explain all the observed uniformities of social
94 behaviour, social organization and social change.’ (p. 39). Thus, the programme theory
95 explains in rather abstract terms how, why, when, where and for whom an intervention
96 is expected to work. It is inspired by policy documents, key informant interviews, local
97 customs or received wisdom, and theories from different scientific disciplines such as
98 psychology, sociology, political science and economics.

99 Subsequently, the realist researcher translates this abstract programme theory
100 into more specific CMOCs adapted to the intervention or phenomenon under study.
101 These specific CMOCs are then further refined or disputed. In the last phase, the
102 information from the refined CMOCs is used to refine the initial middle-range
103 programme theory. The latter is then ready to be used and refined by other evaluations.

104 This creates a cyclical process in which subsequent evaluations build on each other, our
105 knowledge slowly but steadily increases and thorough theory building occurs (see
106 Figure 5.1 in Pawson, 2013).

107 This cyclical process is the focus of Ray Pawson's book *'The Science of*
108 *Evaluation'* (Pawson, 2013) and central to the ResQ approach. In order to streamline
109 this process of theory building, Pawson (2013, pp. 86-111) identifies seven organizing
110 principles. First, programme theories are the unit of analysis. These theories are
111 transferable across interventions in different sectors (e.g. education or health). Second,
112 abstraction is key. Theories are only transferable if they use the same concepts;
113 therefore, a certain level of abstraction is necessary when creating the final programme
114 theory at the middle-range level. Third, there are different classes of interventions and
115 each should have their own 'reusable conceptual platform' that contains a common set
116 of programme theories. Each new evaluation can contribute to or refine these
117 frameworks. Fourth, model building should be central to subsequent evaluations,
118 increasingly refining and 'specifying the conditions in which each programme theory
119 applies' (Pawson, 2013, p. 86). Fifth, such model building should be done through
120 adjudication between rival theories. Sixth, evaluating every part of a theory,
121 incorporating every possible idea or theory, and looking at every single instance of an
122 intervention is impossible. We therefore need to trust some part of the intervention
123 while investigating another part of it. Subsequent evaluation or research cycles will
124 cover the uncovered parts. Seventh, from the selection of the relevant theories, to the
125 conceptual abstraction and the adjudication between rival theories, the evaluator needs
126 to make important judgement calls, 'organized skepticism' (Pawson, 2013, p. 86) is
127 crucial in order to ascertain the necessary rigour.

128 These seven organizing principles are crucial to advancing the theory building
129 within realist evaluation and are equally central to the ResQ approach.

130 **Qualitative Comparative Analysis**

131 As mentioned above, we will combine the realist approach with qualitative comparative
132 analysis (QCA) (Ragin, 2014 [1987]). QCA is a set-theoretic method that uses Boolean
133 algebra to compare and analyse multiple cases in order to determine the necessary and
134 sufficient conditions for a predefined outcome to occur (Ragin, 2014 [1987]; Schneider
135 & Wagemann, 2012). Interestingly, according to several authors, it has the same
136 underlying philosophy as realist evaluation (Gerrits & Verweij, 2013; Rutten, 2021).

137 Being a set-theoretic method means that ‘(a) the data consists of set membership
138 scores; (b) relations between social phenomena are modelled in terms of set relations;
139 and (c) the results point to sufficient and necessary conditions’ (Schneider &
140 Wagemann, 2012, p. 6). Thus, central to these methods is the idea that cases can be
141 assigned to sets, which are groups of cases that share a specific condition. For example,
142 a government run health facility is a member of the set of ‘public health facilities’,
143 while a health facility run by a private organization is not. Each case belongs to an
144 infinite number of sets. For example, the sets of ‘urban facilities’, ‘well-performing
145 facilities’, ‘public hospitals’, etc. Importantly, these different sets can and are being
146 seen in relation to each other. For example, all public hospitals are public health
147 facilities, yet not all public health facilities are hospitals. This means that the set of
148 ‘public health facilities’ is a superset of the set of public hospitals and the latter a subset
149 of the former.

150 As Schneider and Wagemann (2012) highlight, this use of theoretical sets does
151 not add much value unless combined with the concepts of sufficiency and necessity.

152 Simply put, when a condition is causally sufficient/necessary for an outcome to occur,
153 the outcome will always/only occur whenever the condition is present. Indeed, when
154 approaching a set relationship through these concepts, we may start thinking about
155 explanations and causality. For example, if the set of well-financed facilities is a subset
156 of the set of well-performing facilities (i.e. all well-financed facilities are also well-
157 performing facilities), it can be said that, in our data set, being well financed is a
158 sufficient condition for being a well-performing facility. This means that making sure
159 that a facility is well-financed is sufficient for it to be well-performing, however it does
160 not mean that a facility cannot be well-performing without being well-financed. Looked
161 at it the other way around, the set of well-performing facilities can be seen as a superset
162 of the set of well-financed facilities, which means that, in this very limited data set,
163 being a well-performing facility is a necessary condition to be a well-financed facility¹.
164 In other words, a well-financed facility has to be well-performing before it can be well-
165 financed. Importantly, a set-theoretic relationship does not necessarily make a claim
166 about causality (Ragin, 2009).

167 What places QCA apart from other set-theoretic methods is its use of truth tables
168 and the principle of logical minimization (Schneider & Wagemann, 2012). I will discuss
169 these two elements while elaborating the process of QCA..

170 First, QCA is a case-based approach in which cases are looked at as a whole, in
171 contrast to analysing variables delinked from the cases (as is done in statistics) (Ragin,
172 2014 [1987]). Cases are defined by their membership in specific sets of conditions,
173 which are selected based on their role in explaining the outcome of interest according to
174 substantive theories. Data on these conditions and the outcome condition for each of the

¹ This symmetry between sufficiency and necessity is only the case in crisp sets (i.e. either a case is in or out) and the truth table contains no empty rows (see below) and not when fuzzy sets are used (i.e. a case can be partly in and out of a set).

175 cases is calibrated into membership scores. When the membership scores are either 1
176 (condition/outcome is present) or 0 (absent), we call this crisp sets. When the
177 membership scores range between 0 and 1, meaning cases may be partially in and out of
178 a particular set, we call this fuzzy sets. This data matrix with the set membership scores
179 is subsequently transformed into a truth table. A truth table ‘displays the data in a
180 matrix of logically possible configurations of causal conditions’ (Ragin, 2014 [1987], p.
181 xxi). Hence, every row contains a combination of the previously identified conditions
182 and the outcome observed in the cases that display the same combination of conditions.
183 Importantly, whether you are using crisp sets or fuzzy sets, the truth table is always
184 binary.²

185 Second, this truth table is analysed using logical minimization,³ which is ‘a
186 process by which the empirical information is expressed in a more parsimonious yet
187 logically equivalent manner by looking for commonalities and differences among cases
188 that share the same outcome’ (Schneider & Wagemann, 2012, p. 9). More specifically,
189 this process is informed by Boolean logic, in which Boolean addition and multiplication
190 play an important role and are equivalent to the logical operators OR and AND,
191 respectively. Thus, the term ‘ $A + C \rightarrow Z$ ’ means that it is sufficient if A OR B (or both)
192 are present in order for outcome Z to be present; similarly ‘ $A * C \rightarrow Z$ ’ or ‘ $AC \rightarrow Z$ ’
193 means that it is sufficient if A AND C are present in order for outcome Z to be present.
194 The minimization process uses the following logic: ‘If two Boolean expressions differ
195 in only one causal condition yet produce the same outcome, then the causal condition
196 that distinguishes the two expressions can be considered irrelevant and can be removed

² We refer to the standard works on QCA for more details (e.g. Schneider & Wagemann, 2012) on how to go from a data matrix containing fuzzy set-membership scores to a truth table.

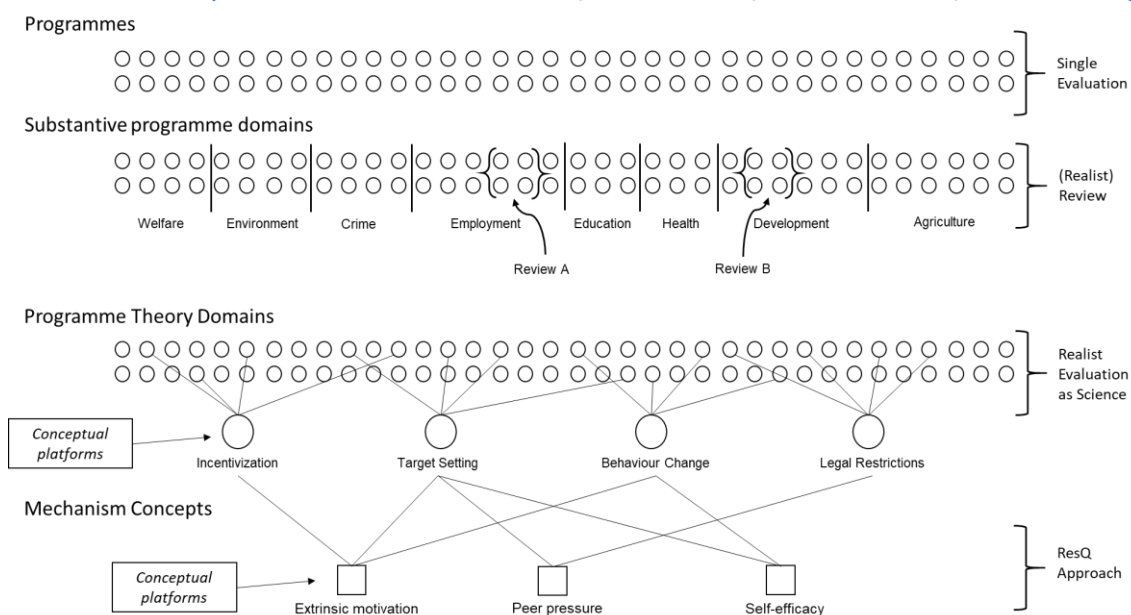
³ Here, we focus on the analysis of sufficient conditions; the analysis of necessary conditions is somewhat different and can be found in Goertz and Starr (2002) and Schneider and Wagemann (2012) among others.

197 to create a simpler, combined expression.’ (Ragin, 2014 [1987], p. 93). The result of
 198 this logical minimization process is a solution term that shows the combination of
 199 factors that are sufficient for the outcome to occur. With this solution term, we can go
 200 back to the case to establish the causal relationship and refine the initial theory.

201 **The ResQ approach**

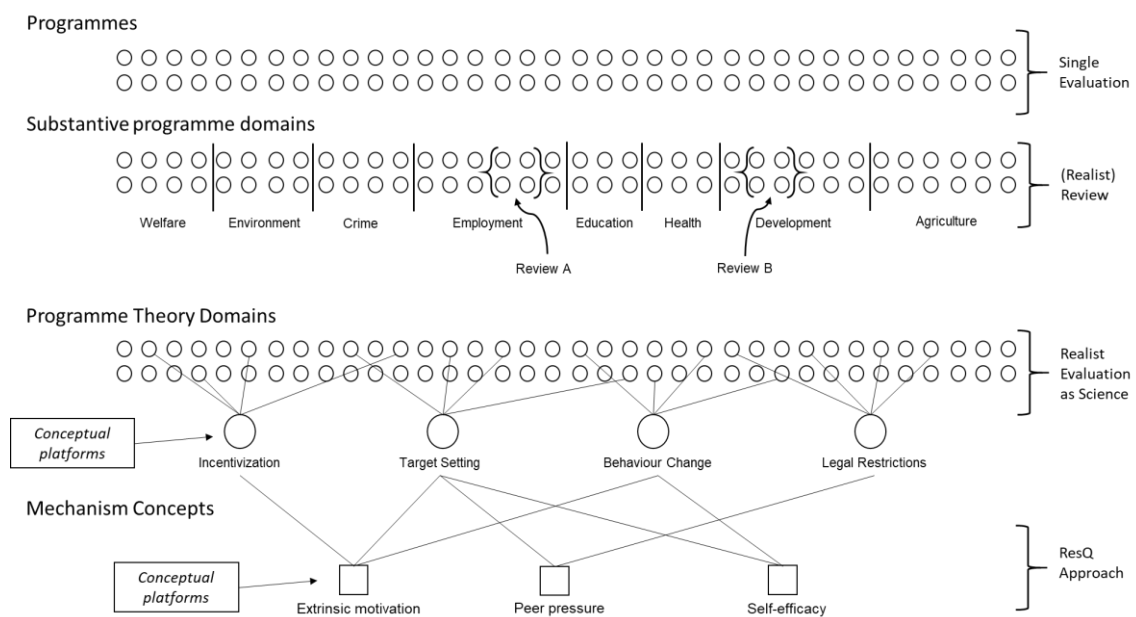
202 The aim of the ResQ approach is to put the plethora of published evaluations and
 203 research to use and build middle-range programme theories. These theories can then be
 204 used to explain why an intervention will possibly work, will possibly not work, has
 205 worked or did not work, and can inform us about useful adaptations to the intervention.

206 *The concept of reusable conceptual platforms (Pawson, 2013) is at the centre of the ResQ approach. It enables us*
 207 *to bridge the domains of evaluation and science and transform evaluation (and research) results into middle-range*
 208 *theories (Marchal et al., 2018; Pawson, 2013).*



209
 210 Figure 1 shows its rationale and how the ResQ approach differs slightly from the
 211 approach put forward by Pawson (2013). As mentioned by Pawson (2013), every circle
 212 represents hundreds of programmes and their evaluations. The second row shows these
 213 programmes per substantive domain. Common systematic reviews focus on similar
 214 programmes within specific domains, however the realist evaluation science approach

215 shifts this focus to programmes with the same underlying programme theory from any
 216 domain (e.g. incentivization programme theory: interventions with vouchers, financial
 217 incentives, favourable loans, etc.) (Pawson, 2013). The ResQ approach takes this a step
 218 further and focuses on the mechanisms constituting these programme theories. Indeed,
 219 the programme theories central to the realist evaluation science approach may consist of
 220 several mechanisms. For example, the programme theory of an incentivization
 221 intervention will include the mechanism of extrinsic motivation but also self-efficacy,
 222 feeling of appreciation, belief updating, etc. In turn, these mechanisms may be present
 223 in very distinct programme theories; for example, extrinsic motivation may be present
 224 in incentivization, target-setting or behaviour change programmes. The ResQ approach
 225 aims to build theories around these mechanisms, that we will call *mechanism concepts*⁴.



226

227 **Figure 1: Focus point of realist evaluation science and the ResQ approach (adapted**
 228 **from Pawson, 2013)**

229 In order to facilitate the use of these *mechanism concepts* as reusable conceptual
 230 platforms, the ResQ approach will mainly make use of established scientific concepts to

⁴ A more extensive discussion of mechanism concepts in the ResQ approach will be discussed in a future paper.

231 increase its ability to connect with scientific theories across the different domains,
232 sectors and disciplines. These ‘tested and robust explanatory theories from the (social,
233 behavioural and policy) sciences [add] crucial insights about mechanisms and contexts’
234 (Leeuw & Donaldson, 2015, p. 472).

235 The mechanism concepts are closely related to the notion of concepts put
236 forward by Goertz (2006). Accordingly, concepts are ‘theories about the fundamental
237 constitutive elements of a phenomenon’ and the analysis of concepts ‘involves
238 ascertaining the constitutive characteristics of a phenomenon that have central causal
239 powers’ (p. 5).⁵ Moreover, Goertz’ (2006) view of concepts is non-causal, meaning that
240 the characteristics we attribute to them are not caused by concepts they refer to or vice
241 versa, but rather constitute them. Likewise, the ResQ approach observes an ontological
242 relationship between the mechanism and the context, in line with the above-mentioned
243 definition of realist mechanisms, which ‘tell us what it is about that system that
244 generates the uniformity [between context and outcome] (Pawson, 2006b, p. 23).
245 Identifying the context elements and their interactions that constitute a mechanism is the
246 essence of the theory-building effort central to the ResQ approach. As Westhorp (2018)
247 puts it: ‘[mechanisms should] involve the description of at least three things: the
248 necessary components of the system, the necessary relationships between those
249 components and the processes (or interactions)’ (p. 53). To this should be added that we
250 not only need to know the necessary but also the sufficient components for a
251 mechanism to possibly ‘fire’, which the ResQ approach identifies through the use of
252 qualitative comparative analysis (QCA).

⁵ Similarly, both Pawson (2013) and Sayer (1992) place a relatively strong emphasis on the need to know the necessary components and scope conditions of the generic conceptual platforms.

253 The ResQ approach is not the first to combine the realist approach with QCA
254 (see for example Befani et al., 2007; Goicolea et al., 2015; Sager & Andereggen, 2012),
255 although it remains uncommon. However, their approach is different from the ResQ
256 approach. The conditions they used in the QCA relate to both the mechanisms and the
257 contextual conditions. This implies that mechanisms and context work hand in hand to
258 generate a certain outcome and may refer to the idea of context as a mediator between
259 the outcome and the mechanism. However, as we saw above, the ResQ approach takes
260 another position on the relationship between mechanism and context.⁶ Indeed,
261 mechanisms lie at a different ontological level than the context and the outcome
262 (Westthorp, 2018; Williams, 2018); or, as mentioned above, mechanisms and context are
263 ontologically, not causally, connected.

264 Given this observation, the ResQ approach takes a somewhat different route,
265 inspired by the multimethod approach put forward by Goertz (2017). The latter
266 differentiates between the causal analysis that is undertaken within the case studies and
267 the cross-case QCA used for generalization or, indeed, theory building. Likewise, as a
268 research synthesis approach, the ResQ approach is based on information from studies in
269 which a causal analysis has been performed, while the QCA technique is used to
270 identify the relevant conditions and extrapolate them to theories (i.e. mechanism
271 concepts).

272 Hence, the triggering of the mechanism (yes or no)⁷ is used as the outcome set
273 and the contextual conditions as condition sets. This means that the ResQ approach
274 performs a ‘C-Mo’ analysis in which the ‘O’ (outcome) is somewhat silent. Indeed, the

⁶ The difference can be related to what Sayer (1992) calls structure and conditions (cf. moderators).

⁷ As will be discussed in a future article, in contrast to Dalkin et al. (2015) I take the position that mechanisms are either triggered or not. My conceptualization of a mechanism states that it is either possessed by a specific system or not and does not leave room for a partial possession. When it seems that they give stronger or weaker outcomes this is only due to the effect of other mechanisms at work at the same time.

275 outcome in common CMO configurations is often general and far down the causal
276 chain, for example improved health-care quality. When studying, for example, extrinsic
277 motivation as a mechanism, a CMOC may indicate that extrinsic motivation in a
278 specific context leads to better health-care quality. However, the absence of improved
279 health-care quality does not mean that the extrinsic motivation mechanism was not
280 triggered. In fact, the lack of proper equipment may hamper the attainment of higher
281 levels of health-care quality.

282 Thus, this focus on more general outcomes blurs our analysis of the mechanism,
283 as it leads to the identification of contextual conditions that are only relevant after the
284 mechanism has been triggered and do not constitute the mechanism⁸. Therefore, it is
285 more fruitful to equate the ‘O’ in the CMO configuration with the ‘proxy equivalents’⁹
286 (Jagosh, 2020) we use to identify the mechanism, for example a higher attendance at a
287 facility or a certain score on a motivation questionnaire. These proxy equivalents are as
288 causally proximate as possible. In practice, our analysis thus becomes a C-M analysis.
289 The decision about which proxy equivalents to use is inspired by theory and should aim
290 at minimizing the possibility that another mechanism can be seen as intervening.

291 When assessing a specific intervention, the relevant mechanism concepts can be
292 seen as the building blocks of a programme theory. The latter helps to explain how the
293 intervention may work in reality and on which aspects the evaluation should focus.
294 Importantly, this is not a predictive model, but it will help us to define what will
295 possibly happen given the current state of knowledge. Knowing what can be expected
296 may help to identify unexpected outcomes, which are valuable learning moments that
297 will further refine the mechanism concepts.

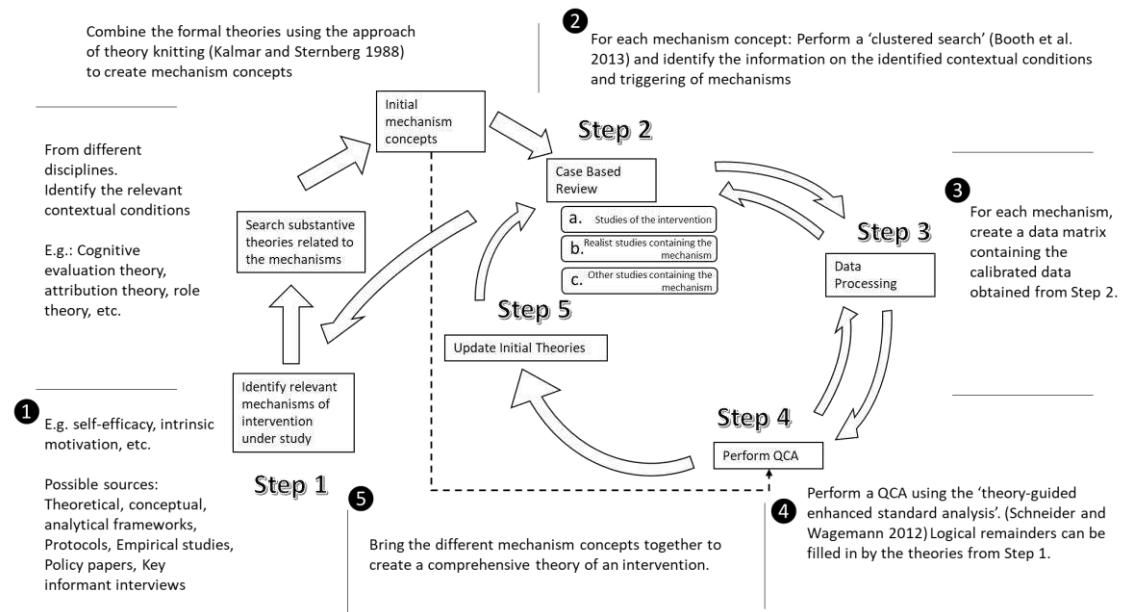
⁸ See note 6.

⁹ Proxy equivalents are small clues that indicate the possible existence of a specific mechanism (Jagosh, 2020).

298 In the long run, a database of mechanism concepts can be constructed, from
299 which evaluators can take the mechanism concepts required and create a programme
300 theory, and after which the mechanism concepts in the database can be further refined.
301 This will help researchers, evaluators and policymakers to maintain an overview of the
302 many findings from the different disciplines and domains in evaluation and research.

303 **Five steps**

304 Having elaborated on the underlying ideas of the ResQ approach, we now turn to a
305 more practical overview of the five steps encompassing its application. Importantly, like
306 both the realist approach and QCA, the ResQ approach is an iterative process in which
307 the researcher can and even should move back and forth between the different steps.
308 This ensures that the most relevant mechanisms are studied, taking into account the
309 latest data and information. This iterative process is summarized in Figure 2, which is
310 followed by an elaborate discussion of the five steps. In the appendix, a short pilot
311 example can be found to illustrate the steps. We use the example of performance-based
312 financing (PBF) in the health care sector of low- and middle-income countries and focus
313 on the mechanism self-efficacy.



314

315 **Figure 2: Overview of the ResQ approach in five steps**

316 **Step 1: Developing the initial mechanism concepts**

317 The ResQ approach starts with the creation of an initial theory, which helps us to focus
 318 our research and attention on those issues, conditions, mechanisms and hypotheses that
 319 are most likely to be relevant (Pawson, 2013). The creation of the theory starts by
 320 identifying the mechanisms that are hypothesized and theorized to be triggered by the
 321 intervention. The main sources for this are the analytical, conceptual and theoretical
 322 frameworks of the intervention or related interventions, research protocols, policy
 323 documents and key informant interviews.

324 Once the possible mechanisms have been identified, we commence the
 325 construction of the 'initial mechanism concepts'. Each of these aims to shed light on the
 326 contextual conditions that correspond to the triggering of each mechanism. To construct
 327 these mechanism concepts, we use the approach of 'theory knitting' (Kalmar &
 328 Sternberg, 1988), by which multiple substantive theories from different disciplines can
 329 be combined. One of the underlying ideas is that rival theories often discuss different
 330 phenomena and thus are not really rivals. The role of the theory developer performing

331 the theory knitting is then to ‘[identify] and effectively [utilize] the common dimensions
332 that underlie the theories to be knit’ (Kalmar & Sternberg, 1988, p. 164), creating a
333 higher order theory (Leeuw & Donaldson, 2015).

334 These substantive theories can be found using a google or library search, or in a
335 substantive theories database created within the framework of the current ResQ study.
336 Importantly, research is a cyclical process and theories that are not found or used during
337 this cycle may be used during another cycle. Therefore, not being able to incorporate or
338 find all the theories is not problematic as long as we remain aware that the final theory
339 and knowledge that we create is only partial (Pawson, 2013; Wong, 2018).

340 Even if many rival theories turn out to be less conflicting, contradictory
341 propositions will remain, and this may be experienced as problematic. However,
342 according to Pawson (2013) ‘[a]djudicating between rival hypotheses is the engine of
343 progress in evaluation science’ (p. 86). Thus, contradictory theories should be seen as
344 an opportunity rather than a problem. To be clear, it is not so much about deciding
345 which theory is wrong, rather ‘[i]t is the realist evaluator’s task [...] to identify and
346 explain the precise circumstances under which each theory holds’ (Pawson, 2013, p. 7).

347 The theory knitting process is aided by the use of a simple truth table which is
348 created for each mechanism, containing the propositions of the substantive theories
349 concerning the contextual factors related to the relevant mechanism. Contradictory
350 theories will be apparent in inconsistent rows and should receive additional attention.

351 For clarity and communication reasons, the researcher may want to create a
352 narrative based on the truth tables for each of the mechanism concepts, or even combine
353 the different mechanism concepts into one programme theory. However, the following
354 steps will be performed at the level of the mechanism concepts.

355 In our example in the appendix I identified 17 mechanisms, including extrinsic
356 motivation, price effect, autonomy and self-efficacy. Here we will focus on self-
357 efficacy. Two theories were used to identify relevant contextual conditions: social
358 cognitive theory (Bandura, 1994) and the theory of planned behaviour (Ajzen, 1991). I
359 identified five relevant contextual conditions: the initial level of self-efficacy, relevant
360 positive experiences, relevant negative experiences, perceived effort put in, and
361 perceived control over the behaviour. Using the truth table and applying the
362 minimization process of QCA gives the following initial mechanism concept which
363 identifies the situations in which the mechanism is triggered (uppercase means the
364 condition is present, lower case means the condition is absent) (see the appendix for
365 more details):

366 POS. EXP.*neg. exp.*EFFORT*CONTROL +
367 INI SE*neg. exp.*effort*CONTROL +
368 INI SE*POS. EXP.*neg. exp.*CONTROL +
369 INI SE*POS. EXP.*neg. exp.*EFFORT +
370 INI SE*pos. exp.*effort +
371 INI SE*pos. exp.*NEG. EXP.

372 ***Step 2: Case-based Search***

373 After having identified the theoretically most important contextual conditions for each
374 of the mechanisms, a case-based review of the empirical literature is performed. The
375 aim is to gather data on these contextual conditions and whether the mechanisms were
376 observed or not. This case-based review may be very intensive and may be performed in
377 several rounds spanning different research projects. Round ‘a.’ focuses on studies
378 covering the intervention under study; during round ‘b.’ the focus is on realist
379 evaluations of any intervention that entails the identified mechanisms; and round ‘c.’

380 looks at non-realist evaluations that study similar interventions, or interventions that
381 will likely trigger similar mechanisms or have similar combinations of contextual
382 conditions. This is also an iterative process in which new searches need to be done
383 whenever new conditions, theories or mechanisms arise in the process (Booth et al.,
384 2020).

385 The data search process will use the technique of ‘clustered search’ (Booth et al.,
386 2013). This is a case-based search method used ‘to identify papers or other research
387 outputs that relate to a single study’ (Booth et al., 2013, p. 4). This will help to gather
388 the necessary data on the contextual conditions, which may not always be reported on in
389 a single paper or document.

390 Finally, we assess the utility and the methodological quality of the information
391 collected for the analysis. In assessing this quality, it is not the study or document as a
392 whole that is taken into account; rather, the quality appraisal concerns whether the
393 information collected for the analysis is supported by the data and the methodology
394 used in that document (Pawson, 2006a; Wong, 2018). Thus, a study may be of bad
395 quality in general, but still provide some interesting and reliable pieces of information
396 for the analysis.

397 In our pilot example, we searched for papers evaluating a PBF intervention and
398 realist evaluations mentioning ‘self-efficacy’ as a mechanism (we did not perform round
399 ‘c.’). Our search shows that self-efficacy is under researched in the PBF literature, yet,
400 despite this, the ResQ approach enables us to learn about this mechanism by including
401 papers from other interventions.

402 ***Step 3: Data processing***

403 Once the relevant papers and documents are gathered, data processing can start. We first

404 need to identify the realist mechanisms that have been observed in the selected studies
405 and documents. While the realist studies of round ‘b.’ will give relatively
406 straightforward information due to the common philosophical underpinnings, the
407 studies from rounds ‘a.’ and ‘c.’ will not report realist mechanisms. The researcher
408 therefore needs to remain flexible and search for ‘proxy equivalents’ of mechanisms
409 (Jagosh, 2020). Such equivalents are not only to be found in the conclusion or the
410 discussion of the studies but should equally be looked for in quotations from interviews,
411 quantitative data and/or observations by the authors (Wong, 2018).

412 It is possible and likely that new mechanisms will be found during the analysis,
413 which means that for that specific mechanism we need to go back to the first step and
414 create a new initial mechanism concept. This shows the iterative nature of the approach.

415 Once mechanisms are found in the documents, we can start collecting data on
416 the contextual conditions that were identified as relevant in Step 1. To do this, we create
417 a data matrix for each mechanism containing information on the contextual conditions.
418 Again, new relevant conditions may be found in the documents and might be added
419 during the process.

420 This is where QCA becomes relevant. The data matrix contains the collected
421 data and the linked calibrated set-membership scores in each of the condition and
422 outcome sets. Although fuzzy sets (membership scores range between 0 and 1) have
423 more detailed information than crisp sets (membership scores are either 0 or 1) and are,
424 therefore, preferred whenever possible (Schneider & Wagemann, 2012), our
425 conceptualization of a mechanism states that it is either possessed by a specific system
426 or not and does not leave room for a partial possession¹⁰. This means that our outcome

¹⁰ As mentioned earlier, this is different from the conceptualization put forward by for example Dalkin et al. (2015) and will be more extensively discussed in a future paper.

427 (i.e. the triggering of a mechanism) has to be a crisp set, implying that our QCA also
428 needs to be a crisp set QCA.

429 Whereas some conditions might be expressed qualitatively (present or not) some
430 will be quantitative data and will need to be calibrated to crisp set-membership scores.
431 To do so, one anchor point is especially important: the cut-off point. This quantitative or
432 qualitative point indicates when a case becomes a member or a non-member of a set
433 (Ragin, 2014 [1987]). Importantly, the distinction between a score of 0 and a score of 1
434 is an inherently qualitative difference and hence needs to make sense within the context
435 of the case and within the theoretical propositions. Therefore, theory and knowledge of
436 the cases guide this decision.

437 The conditions identified in our pilot example are all expressed in a
438 dichotomous, qualitative manner, hence the data matrix consists of 0's and 1's
439 indicating the absence and presence of the condition respectively. Therefore, no
440 calibration was needed. The data matrix can be found in the appendix.

441 ***Step 4: Perform a Qualitative Comparative Analysis (QCA)***

442 Having created the data matrix for each of the mechanisms, we can now start analysing
443 this data using QCA. The data matrix will be transformed into a truth table. A truth
444 table consists of 2^k rows (k being the number of conditions) that depict the various
445 possible combinations of the dichotomized conditions (e.g. high incentives/no high
446 incentives). Based on the information from the data matrix, the final column indicates
447 whether the combination of conditions in that row is indeed a subset of the outcome (in
448 this case, whether the mechanism was present or not). A more elaborate explanation of
449 how to transform a data matrix into a truth table can be found in the standard works on
450 QCA (Ragin, 2009, 2014 [1987]; Schneider & Wagemann, 2012).

451 Now, we perform a minimization on the truth table using Boolean logic to
452 discover which conditions are sufficient for the mechanism to occur. To do this, we use
453 the ‘theory-guided enhanced standard analysis’ proposed by Schneider and Wagemann
454 (2012). This means that theoretical knowledge (called directional expectations) from
455 our first step will be used to fill in the logical remainders (truth table rows that are not
456 filled in because the combination of conditions was not observed in the case studies). If
457 the rows give overly inconsistent information (i.e. not all the cases indicate that the
458 conditions are a subset of the outcome), it can be useful to go back to the cases and
459 determine whether other conditions may have played a role and/or whether the
460 calibration was done incorrectly. However, it can also be due to the contingency of
461 mechanisms. The final result will be a solution term describing the combination of
462 conditions that are sufficient for the mechanisms to be triggered. This can then be
463 interpreted, creating a comprehensive mechanism concept.

464 In our pilot example, one row contradicted the theory which led to an adaptation
465 of the initial mechanism concept. Truth table rows not covered by our cases were filled
466 in using the initial mechanism concept propositions. The following solution term was
467 the outcome of the minimization process performed using the QCA add-in for excel
468 developed by Cronqvist (2019):

469 POS. EXP.*neg. exp.*EFFORT*CONTROL +
470 INI SE*neg. exp.*effort*CONTROL +
471 INI SE*pos. exp.*NEG. EXP.*CONTROL +
472 INI SE*POS. EXP.*neg. exp.*CONTROL +
473 INI SE*POS. EXP.*neg. exp.*EFFORT +
474 INI SE*pos. exp.*effort

475 ***Step 5: Refine initial theories***

476 The final step of the ResQ approach consists of going back to the first step and
477 comparing our final outcome with the initial mechanism concepts. We can refine and
478 update the different substantive middle-range theories used to theory knit the initial
479 mechanism concept and also articulate a refined and updated mechanism concept as a
480 whole. Finally, combining the different building blocks (i.e. the mechanism concepts)
481 will create a comprehensive evidence-based and theory-embedded theory of the
482 intervention that can help to inform future evaluations and research and to explain and
483 resolve possible bottlenecks. Such a comprehensive theory can be presented using a
484 narrative or a diagram such as a causal loop diagram (Tomoaia-Cotisel et al., 2017).
485 Future evaluations and research can subsequently contribute to this by adding more
486 theoretical layers in Step 1, by including more studies from different domains, by
487 empirically testing the mechanism concepts and by doing research on the logical
488 remainders identified in Step 4.

489 While only focusing on one mechanism, our pilot example did give an
490 interesting insight into how PBF can help trigger self-efficacy. Indeed, PBF can have an
491 important impact on the triggering of the self-efficacy mechanism by creating moments
492 for constructive and positive feedback, increasing the effort of the health workers
493 through financial incentives and improving the work environment by increasing the
494 funds of the health facility and incentivizing the facilities. Moreover, it shows that
495 initial self-efficacy plays an important role in triggering the self-efficacy mechanism,
496 which may explain the different outcomes across health workers and may reinforce
497 quality of care inequities between health workers and/or facilities.

498 **Discussion**

499 Having gone through the underlying rationale and the different steps of the ResQ
500 approach, it is time to answer the ‘so what?’-question. Clearly, the approach can be very
501 time consuming and few evaluations will be able to perform such an elaborate task.
502 However, this is not the aim of the ResQ approach, which is, rather, to become a
503 continuous research endeavour to which researchers and evaluators from all around the
504 world can contribute. This will eventually create a database comprising the different
505 mechanism concepts which can be used by evaluators as well as be further refined,
506 effectively creating a public good similar to the Cochrane database. The investment of
507 every individual researcher thus becomes minimal, but the combined effort of these
508 individual researchers will help to make enormous strides in building stronger
509 (programme) theories.

510 This bold ambition does not mean that the ResQ approach aims to replace other
511 methodologies, it is just one tool in our tool box to understand the world around us.
512 Realist evaluation and synthesis and other approaches remain crucial in their own way
513 and in providing the resources (i.e. data) for the ResQ approach. Moreover, the latter
514 does not contradict with Pawson’s proposed approach to realist evaluation science,
515 which focuses on program theories at the level of intervention families (Pawson, 2013).
516 In fact it is quite compatible as the mechanism concepts can inform Pawson’s program
517 theories which in turn help to inform us about the interaction between the mechanisms.

518 In spite of the expected merits of the approach it also entails certain limitations and
519 possible criticisms from other realist evaluators. For instance, the ResQ approach is
520 more inclined towards science than it is towards evaluation when it comes to the
521 mechanisms. It prioritizes established scientific concepts that can connect with a

522 relatively large pool of empirical studies and scientific theories. Seasoned realist
523 evaluators may suggest that this may lead to overlooking the main intricacies of a
524 mechanism observed in a realist evaluation. This is the well-known trade-off between
525 breadth and depth, which is part of every research synthesis method. However, the loss
526 of depth is strongly off-set by the large increase in breadth. As the mechanism concepts
527 become increasingly detailed and refined, we will come closer and closer to
528 incorporating these intricacies.

529 Another point of contention may be the focus on the triggering of the mechanism and
530 not on a specific outcome as such. However, the realist approach is built up around
531 CMO configurations in which the outcome plays an important role. We have already
532 discussed our reasons for focusing on the mechanism; however, this leaves open the
533 question of how to link the mechanisms to specific outcomes. For this, we point to the
534 definition of the mechanism, which entails the expected outcome when made
535 sufficiently precise. For example, extrinsic motivation might be defined as ‘motivation
536 due to an extrinsic reward’. Based on this definition, we cannot indicate what the
537 outcome will be when triggered. However, the following definition already tells us
538 more: ‘motivation to perform a specific task due to an external reward linked to that
539 task’. This implies that only the tasks related to receiving a reward can be expected to
540 improve, unless other mechanisms, such as intrinsic motivation, are also triggered.
541 Thus, when tailoring the mechanism concepts to a specific intervention, i.e. linking
542 them to each other, the effect on the outcome of interest will become clear.

543 The ResQ approach needs an important disclaimer: theories coming out of the ResQ
544 approach are not to be used as predictive models. They are explanatory and can help to
545 focus our attention on possible bottlenecks. Every mechanism concept is based on the

546 abundant but still limited knowledge we have at our disposal and is therefore by
547 definition flawed. Moreover, in the realist view of the social world, we may never
548 achieve full knowledge due to the volition of human actors, their unpredictability and
549 their distinct rationalities (Bhaskar, 2008 [1975]; Pawson, 2013).

550 **Conclusion**

551 In this paper I have introduced a novel methodology that aims to further the ambition
552 expressed by Pawson (2013) to create a true evaluation cycle in which consecutive
553 evaluations build on each other's findings. I propose to combine Pawson's (2013) realist
554 evaluation science with qualitative comparative analysis (Ragin, 2014 [1987]) to
555 develop mechanism concepts which express theories that indicate in which
556 circumstances certain mechanisms may be triggered. Using a pilot example, I have
557 shown how the ResQ approach works out in practice (see appendix). More work needs
558 to be done to analyse more papers and include them in the ResQ mechanism database.

559 **References**

- 560 Abner, G. B., Kim, S. Y., & Perry, J. L. (2017). Building Evidence for Public Human
561 Resource Management: Using Middle Range Theory to Link Theory and Data.
562 *Review of Public Personnel Administration*, 37(2), 139-159.
563 doi:10.1177/0734371x17697248
- 564 Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human*
565 *Decision Processes*, 50(2), 179-211. doi:[https://doi.org/10.1016/0749-](https://doi.org/10.1016/0749-5978(91)90020-T)
566 [5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- 567 Astbury, B., & Leeuw, F. L. (2010). Unpacking Black Boxes: Mechanisms and Theory
568 Building in Evaluation. *American Journal of Evaluation*, 31(3), 363-381.
569 doi:10.1177/1098214010371972
- 570 Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of*
571 *human behavior* (Vol. 4, pp. 71-81). New York: Academic Press.
- 572 Befani, B., Ledermann, S., & Sager, F. (2007). Realistic Evaluation and
573 QCA: Conceptual Parallels and an Empirical Application. *Evaluation*, 13(2),
574 171-192. doi:10.1177/1356389007075222
- 575 Bhaskar, R. (2008 [1975]). *A realist theory of science*. London: Verso books.
- 576 Booth, A., Briscoe, S., & Wright, J. M. (2020). The "realist search": A systematic
577 scoping review of current practice and reporting. *Research Synthesis Methods*,
578 11(1), 14-35. doi:<https://doi.org/10.1002/jrsm.1386>

- 579 Booth, A., Harris, J., Croot, E., Springett, J., Campbell, F., & Wilkins, E. (2013).
580 Towards a methodology for cluster searching to provide conceptual and
581 contextual “richness” for systematic reviews of complex interventions: case
582 study (CLUSTER). *BMC Medical Research Methodology*, *13*(1), 118.
583 doi:10.1186/1471-2288-13-118
- 584 Byrne, D. (2018). Researching complex large-scale nested interventions. In N. Emmel,
585 J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist
586 research* (pp. 91-106). London: Sage.
- 587 Cronqvist, L. (2019). QCA Add-In [Version 1.1]. Retrieved from [https://www.qca-](https://www.qca-addin.net)
588 [addin.net](https://www.qca-addin.net)
- 589 Crossref. (2019). Cross ref Annual Report & Fact File 2018-19. Retrieved from
590 <https://doi.org/10.13003/y8ygmw5>
- 591 Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015).
592 What’s in a mechanism? Development of a key concept in realist evaluation.
593 *Implementation Science : IS*, *10*, 49. doi:10.1186/s13012-015-0237-x
- 594 Emmel, N. (2021). Post-disciplinary realism. *International Journal of Social Research
595 Methodology*, *24*(1), 95-108. doi:10.1080/13645579.2020.1803526
- 596 Gerrits, L., & Verweij, S. (2013). Critical Realism as a Meta-Framework for
597 Understanding the Relationships between Complexity and Qualitative
598 Comparative Analysis. *Journal of Critical Realism*, *12*(2), 166-182.
599 doi:10.1179/rea.12.2.p663527490513071
- 600 Goertz, G. (2006). *Social Science Concepts: A User's Guide*. Princeton: Princeton
601 University Press.
- 602 Goertz, G. (2017). *Multimethod Research, Causal Mechanisms, and Case Studies: An
603 Integrated Approach*. Princeton: Princeton University Press.
- 604 Goertz, G., & Starr, H. (Eds.). (2002). *Necessary Conditions: Theory, Methodology, and
605 Applications*: Rowman & Littlefield.
- 606 Goicolea, I., Vives-Cases, C., Hurtig, A.-K., Marchal, B., Briones-Vozmediano, E.,
607 Otero-García, L., . . . San Sebastian, M. (2015). Mechanisms that Trigger a
608 Good Health-Care Response to Intimate Partner Violence in Spain. Combining
609 Realist Evaluation and Qualitative Comparative Analysis Approaches. *PLoS
610 One*, *10*(8), e0135167. doi:10.1371/journal.pone.0135167
- 611 Jagosh, J. (2020). Retroductive theorizing in Pawson and Tilley's applied scientific
612 realism. *Journal of Critical Realism*, *19*(2), 121-130.
613 doi:10.1080/14767430.2020.1723301
- 614 Kalmar, D. A., & Sternberg, R. J. (1988). Theory knitting: An integrative approach to
615 theory development. *Philosophical Psychology*, *1*(2), 153-170.
616 doi:10.1080/09515088808572934
- 617 Leeuw, F. L., & Donaldson, S. I. (2015). Theory in evaluation: Reducing confusion and
618 encouraging debate. *Evaluation*, *21*(4), 467-480.
619 doi:10.1177/1356389015607712
- 620 Marchal, B., Kegels, G., & Van Belle, S. (2018). Theory and realist methods. In N.
621 Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing
622 realist research* (pp. 79-89). London: Sage.
- 623 Merton, R. K. (1968). *Social theory and social structure*. New York: Free Press.
- 624 Pawson, R. (2006a). Digging for Nuggets: How ‘Bad’ Research Can Yield ‘Good’
625 Evidence. *International Journal of Social Research Methodology*, *9*(2), 127-142.
626 doi:10.1080/13645570600595314
- 627 Pawson, R. (2006b). *Evidence-Based Policy: A Realist Perspective*. London: SAGE
628 Publications.

- 629 Pawson, R. (2013). *The science of evaluation: A realist manifesto*. London: Sage
630 publications.
- 631 Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). Realist synthesis: an
632 introduction. *ESRC Research Methods Programme Working Paper, 2/2004*.
- 633 Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- 634 Perry, J. L. (2010). A Strategic Agenda for Public Human Resource Management
635 Research. *Review of Public Personnel Administration, 30*(1), 20-43.
636 doi:10.1177/0734371x09351821
- 637 Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago:
638 University of Chicago Press.
- 639 Ragin, C. C. (2014 [1987]). *The Comparative Method: Moving Beyond Qualitative and*
640 *Quantitative Strategies*. Berkeley, CA: University of California Press.
- 641 Rutten, R. (2021). Uncertainty, possibility and causal power in QCA. *Sociological*
642 *Methods & Research*. doi:Not yet available
- 643 Sager, F., & Andereggen, C. (2012). Dealing With Complex Causality in Realist
644 Synthesis: The Promise of Qualitative Comparative Analysis. *American Journal*
645 *of Evaluation, 33*(1), 60-78. doi:10.1177/1098214011411574
- 646 Sayer, A. (1992). *Method in Social Science: A Realist Approach*. London, New York:
647 Routledge.
- 648 Sayer, A. (2000). *Realism and Social Science*. London: SAGE Publications.
- 649 Schneider, C. Q., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social*
650 *Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge
651 University Press.
- 652 Tomoaia-Cotisel, A., Kim, H., Allen, S. D., & Blanchet, K. (2017). Causal loop
653 diagrams: a tool for visualizing emergent system behaviour. In D. de Savigny,
654 K. Blanchet, & T. Adam (Eds.), *Applied systems thinking for health systems*
655 *research: A methodological handbook* (pp. 97 - 114). London: Open University
656 Press.
- 657 Westhorp, G. (2018). Understanding mechanisms in realist evaluation and research. In
658 N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.),
659 *Doing realist research* (pp. 41-57). London: Sage Publications.
- 660 Williams, M. (2018). Making up mechanisms in realist research. In N. Emmel, J.
661 Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist*
662 *research* (pp. 25-40). London: Sage.
- 663 Wong, G. (2018). Data gathering in realist reviews: Looking for needles in haystacks. In
664 N. Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.),
665 *Doing realist research* (pp. 131-145). London: Sage.

666

667

669 **Pilot example of the ResQ approach:**

670 **Performance-based financing in the health care sector of low- and middle-income**
 671 **countries: the mechanism self-efficacy**

672 In this appendix I give a short example of how the ResQ approach can be implemented. I do not
 673 give a full example, as this would lead to a full second paper in the appendix. Instead, I show the
 674 workings of the approach by showing a pilot example in which a limited number of papers was
 675 used to pilot test the approach. The intervention under study is called performance-based
 676 financing, which is mainly prevalent in the health care sector of low- and middle-income
 677 countries. For this pilot I focus on one specific mechanism, namely ‘self-efficacy’.

678 Performance-based financing in the health care sector of low- and middle-income countries is
 679 more than just financial incentives, it is a complex intervention with many components
 680 contributing to the final outcomes. It has been defined as follows: “performance-based financing
 681 is a supply-side reform package that is guided towards improved performance (defined as
 682 increased predefined services and improved quality measures) by using performance-based
 683 financial incentives for health providers (facilities and/or workers) through internal contracting
 684 and strengthening this with most or all of the following elements: a separation of functions
 685 (purchaser, provider, verifier), (spending) autonomy for the health facilities, strict monitoring and
 686 verification of services, community involvement, result-based planning and accountability
 687 arrangements.” (Renmans et al., 2017)

688 These different components lead to different mechanisms being triggered. In what follows I show
 689 the pilot test of the developed methodology using self-efficacy as an example.

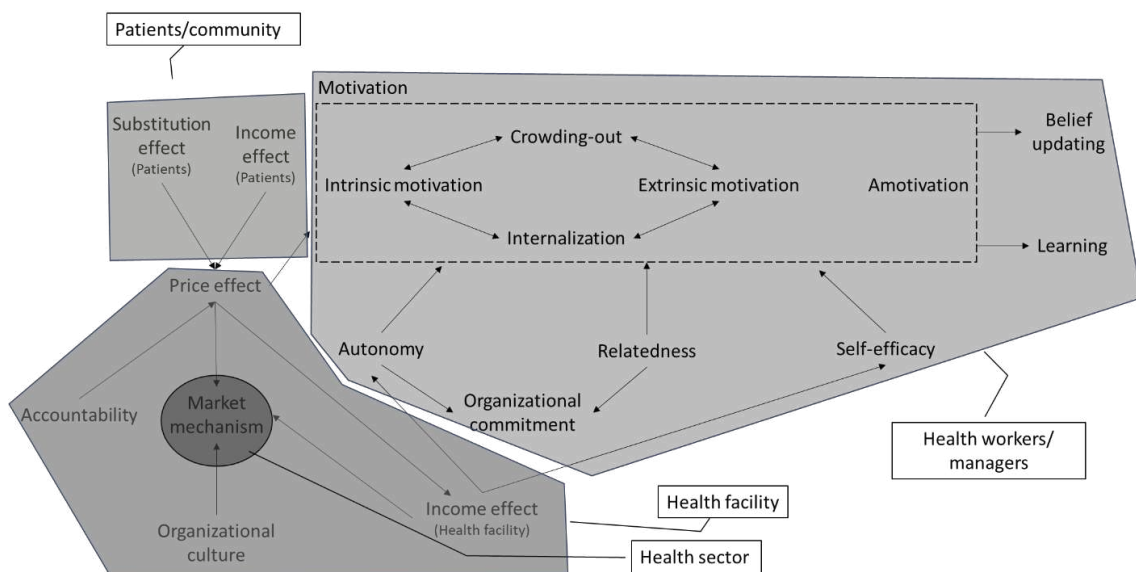
690 *Step 1: Developing the initial mechanism concepts*

- 691 - Identify the relevant mechanisms

692

693 I selected six relevant documents: three protocols (Borghi et al., 2018; Nimpagaritse et al.,
 694 2016; Ridde et al., 2014), a theoretical discussion of the effect of PBF on motivation
 695 (Lohmann et al., 2016), a comprehensive PBF toolkit published by the World Bank (Fritsche
 696 et al., 2014) and a conceptual framework developed by the Health Results Innovation Trust
 697 Fund of the World Bank (HRITF, 2015). Figure 1 shows the different mechanisms and how
 698 they are possibly linked to each other.

699 *Figure 3: Relevant mechanisms triggered by PBF*



700

701 - Theory knitting the initial mechanism concept using a truth table, example: SELF-
 702 EFFICACY

703
 704 To develop the initial mechanism concept and perform the theory-knitting I selected two
 705 theories: social cognitive theory (Bandura, 1994) and the theory of planned behavior (Ajzen,
 706 1991). Table 1 shows the different relevant conditions identified in the literature on these
 707 two theories. In the column ‘condition 2’ I made abstraction of the conditions from the
 708 column ‘condition 1’. This led to the following relevant conditions: relevant positive
 709 experience, relevant negative experience, effort, initial self-efficacy, and perceived control.

710 *Table 1: Deriving the relevant contextual conditions from the theories*

Theory	Condition1		Condition 2
Social Cognitive Theory	Own positive experience	-->	Relevant positive experience
	Sustained effort	-->	Effort
	Prior self-efficacy	-->	Initial self-efficacy
	Role model with positive experience	-->	Relevant positive experience
	Similarity w/ role model	-->	Relevant positive experience
	Acquisition of better means	-->	Control
	Social persuasion of SE	-->	Relevant positive experience
	Focus on self-improvement	-->	Relevant positive experience
	Co-operative learning structure	-->	Relevant positive experience
	Conducive situation	-->	Control
	Fatigue, stress, anxiety	-->	Relevant negative experience
	Collective perception of SE	-->	Not retained
	Theory of Planned Behavior	Required resources	-->
Few and/or manageable obstacles		-->	Control
Belief about resources		-->	Control
Belief about opportunities		-->	Control
positive behavior/exp. of acquaintances and friends		-->	Relevant positive experience
Positive past experiences		-->	Relevant positive experience

711
 712 Table 2 shows the data matrix created based on the two theories. In the aforementioned
 713 papers we looked for the configurations that the authors put forward that would lead to the
 714 triggering or not of self-efficacy. The points (.) indicate that the theory was not explicit about
 715 the value of this condition and, hence, can be either.

716 *Table 2: Data matrix of propositions in theories*

Row	Theory	Ini SE	Pos. Exp.	Neg. Exp.	Effort	Control	Outcome
A	SCT	.	1	0	1	1	1
B	SCT	.	1	0	.	.	1
C	SCT	.	1	0	0	.	0
D	SCT	.	1	.	.	1	1
E	SCT	.	1	.	.	0	0
F	SCT	.	1	.	.	.	1
G	SCT	0	0	1	0	.	0

H	SCT	.	0	0	1	0	0
I	SCT	0	.	1	.	.	0
J	SCT	1	.	1	.	.	1
K	TPB	.	1	.	.	1	1
L	TPB	.	1	.	.	0	0

717

718 Table 3 shows the truth table based on the data matrix from Table 2. The points (.) in the
719 data matrix were interpreted as follows: either present or absent. This means that each row
720 from the data matrix can be used for different rows in the truth table (for example row 'B'
721 from the data matrix). This also leads to contradicting rows for which I needed to decide
722 which one to keep based on the theories at hand. The reasons for certain choices can be found
723 in Table 4. I also deleted some rows because they were impossible, for example when there
724 is both a positive and a negative experience (the assumption is that one always prevails over
725 the other or that there has not been an experience yet).

726

Table 3: Truth table according to the theories

Row	Ini SE	Pos. Exp.	Neg. Exp.	Effort	Control	1	0	Outcome
1	1	1	1	1	1	DFJKL		
2	1	1	1	1	0	FJK	E	
3	1	1	1	0	1	DFJKL		
4	1	1	1	0	0	FJK	E	
5	1	1	0	1	1	A BDFK		1
6	1	1	0	1	0	BF	EL	1
7	1	1	0	0	1	BDFK	C	1
8	1	1	0	0	0	BF	CEL	0
9	1	0	1	1	1	J		1
10	1	0	1	1	0	J		1
11	1	0	1	0	1	J		1
12	1	0	1	0	0	J		1
13	1	0	0	1	1			
14	1	0	0	1	0		H	
15	1	0	0	0	1			1
16	1	0	0	0	0			1
17	0	1	1	1	1	DFKL	I	
18	0	1	1	1	0	FK	EI	
19	0	1	1	0	1	DFKL	I	
20	0	1	1	0	0	FK	EI	
21	0	1	0	1	1	ABDFK		1
22	0	1	0	1	0	BF	EK	0
23	0	1	0	0	1	BDFK	C	0
24	0	1	0	0	0	BF	CEK	0
25	0	0	1	1	1		I	0
26	0	0	1	1	0		I	0
27	0	0	1	0	1		GI	0
28	0	0	1	0	0		GI	0

29	0	0	0	1	1			
30	0	0	0	1	0		H	
31	0	0	0	0	1			0
32	0	0	0	0	0			0

727

728 Table 4: Explanation for the choices of the outcomes when theories conflicted or the theories did not predict anything

- 1 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 2 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 3 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 4 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 5
- 6 I say '1' because 'pos. Exp.' with 'effort' strengthens the already existing 'Ini. SE', the lack of 'control' will not affect the SE as the agent will blame the environment instead of themselves. Which means the SE remains intact.
- 7 Although the lack of 'effort' minimizes the effect of the 'Pos. Exp.' the fact that there is already 'Ini. SE' makes that it still strengthens SE
- 8 Although there is 'initial SE', the lack of effort used to achieve the 'Pos Exp' makes the SE less strong, if we include that the control is limited this will lead to low SE as small setbacks will be deteriorate the SE quickly.
- 9
- 10
- 11
- 12
- 13 Cases that combine 'effort' = 1 with 'Pos. Exp.' = 0 and 'Neg. Exp.' = 0 have been left out because impossible as effort relates to the experiences
- 14 Cases that combine 'effort' = 1 with 'Pos. Exp.' = 0 and 'Neg. Exp.' = 0 have been left out because impossible as effort relates to the experiences
- 15 No cases in the theories, however, because nothing is affecting the initial SE negatively it receives a 1
- 16 No cases in the theories, however, because nothing is affecting the initial SE negatively it receives a 1
- 17 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 18 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 19 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 20 Cases that combine 'Pos. Exp.' With 'Neg. Exp.' have been left out because impossible
- 21
- 22 There is no perceived control over the behavior.
- 23 Lack of effort makes that the SE is fragile.
- 24 Lack of effort means that the SE is fragile and a lack of control means they do not see the appropriate environment as present.
- 25
- 26
- 27
- 28
- 29 Cases that combine 'effort' = 1 with 'Pos. Exp.' = 0 and 'Neg. Exp.' = 0 have been left out because impossible as effort relates to the experiences
- 30 Cases that combine 'effort' = 1 with 'Pos. Exp.' = 0 and 'Neg. Exp.' = 0 have been left out because impossible as effort relates to the experiences
- 31 No cases in the theory, but because no initial SE, SE remains absent, the idea of control over the behavior is insufficient to lead to self-efficacy.
- 32 No cases in the theory, but because no initial SE, SE remains absent.

729

730 - Applying the minimization method of QCA to come to a solution term that depicts the
731 initial mechanism concept

732 After having created this truth table I perform a minimization on this table using the QCA
733 add-in for excel developed by Cronqvist (2019). This gives the following solution term
734 (conditions in uppercase and underlined are present) ¹¹:

735 POS. EXP.*neg. exp.*EFFORT*CONTROL +

736 INI SE*neg. exp.*effort*CONTROL +

737 INI SE*POS. EXP.*neg. exp.*CONTROL +

738 INI SE*POS. EXP.*neg. exp.*EFFORT +

739 INI SE*pos. exp.*effort +

740 INI SE*pos. exp.*NEG. EXP.

741 This solution term shows the conditions under which self-efficacy is being triggered according to
742 the theories used in this pilot example. This shows the importance of the initial level of self-
743 efficacy, on the one hand, and, on the other hand, the importance of positive experiences, effort
744 related to those experiences and the perception of control whenever no self-efficacy is present.

745 *Step 2 Case-based search*

746 - Search for papers on performance-based financing (Round a.)

747

748 Based on an update of the search performed for an earlier systematic review, I collected 129
749 papers on performance-based financing in LMIC. I searched for self-efficacy in these papers
750 (the actual analysis is more systematic going through each paper and identifying each
751 mechanism) and focused on one specific PBF intervention in Malawi (RBF4MNH). The
752 following papers related to studies on this intervention and were used to find information on
753 the conditions that trigger self-efficacy as a mechanism:

754

- 755 1. Lohmann, J., Muula, A. S., Houlfort, N., & De Allegri, M. (2018). How does
756 performance-based financing affect health workers' intrinsic motivation? A Self-
757 Determination Theory-based mixed-methods study in Malawi. *Social Science &
758 Medicine*, 208, 1-8.
- 759 2. Lohmann, J., Wilhelm, D., Kambala, C., Brenner, S., Muula, A. S., & De Allegri,
760 M. (2018). 'The money can be a motivator, to me a little, but mostly PBF just helps
761 me to do better in my job.' An exploration of the motivational mechanisms of
762 performance-based financing for health workers in Malawi. *Health Policy Plan*,
763 33(2), 183-191. doi:10.1093/heapol/czx156
- 764 3. Brenner, S., Wilhelm, D., Lohmann, J., Kambala, C., Chinkhumba, J., Muula, A.
765 S., & De Allegri, M. (2017). Implementation research to improve quality of
766 maternal and newborn health care, Malawi. *Bull World Health Organ*, 95(7), 491-
767 502. doi:10.2471/blt.16.178202
- 768 4. Chinkhumba, J., De Allegri, M., Mazalale, J., Brenner, S., Mathanga, D., Muula,
769 A. S., & Robberstad, B. (2017). Household costs and time to seek care for
770 pregnancy related complications: The role of results-based financing. *PLoS One*,

¹¹ To be correct, the analysis showed two different solutions that are logically equivalent. This is because during the minimization process, combinations can be combined in different ways. However, here I include all the prime implicants involved in those two solution terms. This means that one prime implicant in this solution term is logically redundant. However, since parsimoniousness is not the ultimate objective here, this is not problematic.

12(9), e0182326. doi:10.1371/journal.pone.0182326

5. Kambala, C., Lohmann, J., Mazalale, J., Brenner, S., Sarker, M., Muula, A. S., & De Allegri, M. (2017). Perceptions of quality across the maternal care continuum in the context of a health financing intervention: Evidence from a mixed methods study in rural Malawi. *BMC Health Serv Res*, 17(1), 392. doi:10.1186/s12913-017-2329-6
6. Wilhelm, D. J., Brenner, S., Muula, A. S., & De Allegri, M. (2016). A qualitative study assessing the acceptability and adoption of implementing a results based financing intervention to improve maternal and neonatal health in Malawi. *BMC Health Serv Res*, 16(1), 398. doi:10.1186/s12913-016-1652-7
7. Chinkhumba, J., De Allegri, M., Brenner, S., Muula, A., & Robberstad, B. (2020). The cost-effectiveness of using results-based financing to reduce maternal and perinatal mortality in Malawi. *BMJ Glob Health*, 5(5). doi:10.1136/bmjgh-2019-002260

- Search for realist evaluation papers that include self-efficacy as a mechanism (Round b.)

For this pilot example, I relied on an ongoing scoping review of realist evaluation studies and I again looked for self-efficacy mentioned in these papers. I selected the following papers:

1. Abejirinde, I. O., Zweekhorst, M., Bardaji, A., Abugnaba-Abanga, R., Apentibadek, N., De Brouwere, V., . . . Marchal, B. Unveiling the Black Box of Diagnostic and Clinical Decision Support Systems for Antenatal Care: Realist Evaluation. *JMIR MHealth and UHealth*, 6(12), e11468.
2. Krishnaratne, S., Hamon, J. K., Hoyt, J., Chantler, T., Landegger, J., Spilotros, N., . . . Webster, J. (2021). What mechanisms drive uptake of family planning when integrated with childhood immunisation in Ethiopia? A realist evaluation. *BMC Public Health*, 21(1), 99.
3. Mukumbang, F. C., van Wyk, B., Van Belle, S., & Marchal, B. (2019). 'At this [adherence] club, we are a family now': A realist theory-testing case study of the antiretroviral treatment adherence club, South Africa. *Southern African Journal of HIV Medicine*, 20(1), 922.
4. Vareilles, G., Marchal, B., Kane, S., Petric, T., Pictet, G., & Pommier, J. (2015). Understanding the motivation and performance of community health volunteers involved in the delivery of health programmes in Kampala, Uganda: a realist evaluation. *BMJ Open*, 5(11), e008614.
5. Lefroy, J., Yardley, S., Kinston, R., Gay, S., McBain, S., & McKinley, R. (2017). Qualitative research using realist evaluation to explain preparedness for doctors' memorable 'firsts'. *Medical Education*, 51(10), 1037-1048.
6. Darlington, E. J., Violon, N., & Jourdan, D. (2018). Implementation of health promotion programmes in schools: an approach to understand the influence of contextual factors on the process? *BMC Public Health*, 18(1), 163.

These papers cover a variety of countries and interventions: Bliss4Midwives in Ghana, family planning intervention in Ethiopia, anti-retroviral adherence club in South Africa, Community Health Volunteers in Uganda, training of doctors (apprenticeship) in UK, and health promotion in schools in France.

Step 3 Data Processing

- Identify mechanisms in cases

823
824
825
826
827
828

Given that this is a pilot test in which I only focus on one mechanism I did not systematically go through each paper and identify every mechanism. However, this is being done in the full approach.

- Collect data from the sources

829
830
831
832
833

Given that all the contextual conditions were qualitative and dichotomous no calibration was needed. I expect this to be a quite prevalent situation in the future. The secondary sources already indicated whether the condition was present or not. Table 5 shows the data matrix for the different cases.

Table 5: Data matrix of cases included in the pilot

Case code	Ini SE	Pos. Exp.	Neg. Exp.	Effort	Control	Outcome
Sef.C.01.Mal	.	1	0	1	1	1
Sef.C.01.Mal	.	0	1	1	0	0
Sef.C.01.Mal	1	0	1	1	1	1
Sef.C.01.Mal	0	0	1	1	1	0
Sef.C.01.Mal	0	1	0	1	0	0
Sef.C.02.Gha	0	1	0	0	1	0
Sef.C.03.Eth	0	1	0	1	1	1
Sef.C.03.Eth	0	0	0	0	1	0
Sef.C.04.Saf	.	1	0	1	1	1
Sef.C.05.Uga	0	1	0	1	1	1
Sef.C.06.UK	0	1	0	1	1	1
Sef.C.06.UK	0	0	1	0	1	0
Sef.C.07.Fra	1	1	0	1	1	1
Sef.C.07.Fra	0	0	1	1	0	0

834

835

Step 4 Perform a QCA

836
837
838
839
840
841
842
843
844

- Transform the data matrix into a truth table

Table 6 shows the truth table for the mechanism self-efficacy after including the empirical cases. The columns '1' and '0' indicate how many cases with that combination of conditions showed that the mechanism was present or absent, respectively. The column 'theory' shows what the theory said about the expected outcome. The barred rows are left out of the analysis for the same reason as mentioned above in Table 4. We see that row '10' conflicts with the theory. Therefore, the initial theoretical truth table and the initial mechanism concept needs to be adapted.

845 Table 6: Truth table after analysis of the empirical cases

Row	Ini SE	Pos. Exp.	Neg. Exp.	Effort	Contro l	1	0	Theory	Cases
1	1	1	1	1	1				
2	1	1	1	1	0				
3	1	1	1	0	1				
4	1	1	1	0	0				
5	1	1	0	1	1	3	1	1	Sef.C.01.Mal, Sef.C04.Saf, Sef.C.07.Fra
6	1	1	0	1	0		1	1	
7	1	1	0	0	1		1	1	
8	1	1	0	0	0		0	0	
9	1	0	1	1	1	1	1	1	Sef.C.01.Mal,
10	1	0	1	1	0		1	1	Sef.C.01.Mal,
11	1	0	1	0	1		1	1	
12	1	0	1	0	0		1	1	
13	1	0	0	1	1				
14	1	0	0	1	0				
15	1	0	0	0	1		1	1	
16	1	0	0	0	0		1	1	
17	0	1	1	1	1				
18	0	1	1	1	0				
19	0	1	1	0	1				
20	0	1	1	0	0				
21	0	1	0	1	1	5	1	1	Sef.C.01.Mal, Sef.C.03.Eth, Sef.C.04.Saf, Sef.C.05.Uga, Sef.C.06.UK,
22	0	1	0	1	0		1	0	Sef.C.01.Mal,
23	0	1	0	0	1		1	0	Sef.C.02.Gha
24	0	1	0	0	0			0	
25	0	0	1	1	1		1	0	Sef.C.01.Mal,
26	0	0	1	1	0		2	0	Sef.C.01.Mal, Sef.C.07.Fra
27	0	0	1	0	1		1	0	Sef.C.06.UK
28	0	0	1	0	0			0	
29	0	0	0	1	1				
30	0	0	0	1	0				
31	0	0	0	0	1		1	0	Sef.C.03.Eth
32	0	0	0	0	0			0	

846

847

848 - Perform a minimization on the truth table

849

850

851

As with the initial theoretical truth table I perform a minimization on this table using the QCA add-in for excel developed by Cronqvist (2019). This gives the following solution term (conditions in uppercase and underlined>

¹² To be correct, the analysis showed two different solutions that are logically equivalent. This is because during the minimization process, combinations can be combined in different ways. However, here I include all the prime implicants involved in those

852 POS. EXP.*neg. exp.*EFFORT*CONTROL +
853 INI SE*neg. exp.*effort*CONTROL +
854 INI SE*pos. exp.*NEG. EXP.*CONTROL +
855 INI SE*POS. EXP.*neg. exp.*CONTROL +
856 INI SE*POS. EXP.*neg. exp.*EFFORT +
857 INI SE*pos. exp.*effort

858 *Step 5 Refine initial mechanism concepts*

859 - Refine the initial theories

860
861 Table 7 shows a comparison of the solution term as a result of the theoretical truth table with
862 the solution term after introducing the empirical data. Because the truth table only differed
863 in one instance, there is no big difference in the two solution terms. This is logical as we do
864 not expect the theories to be vastly in contradiction with reality. The refinement to the
865 mechanism concept emphasizes the importance of perceived control of the behavior.

866 *Table 7: Comparison between the initial and refined mechanism concept*

Initial mechanism concept		Refined mechanism concept
POS. EXP.*neg. exp.*EFFORT*CONTROL +	=	POS. EXP.*neg. exp.*EFFORT*CONTROL +
INI SE*neg. exp.*effort*CONTROL +	=	INI SE*neg. exp.*effort*CONTROL +
INI SE*POS. EXP.*neg. exp.*CONTROL +	=	INI SE*POS. EXP.*neg. exp.*CONTROL +
INI SE*POS. EXP.*neg. exp.*EFFORT+	=	INI SE*POS. EXP.*neg. exp.*EFFORT +
INI SE*pos. exp.*effort +	=	INI SE*pos. exp.*effort +
INI SE*pos. exp.*NEG. EXP.	/	INI SE*pos. exp.*NEG. EXP.*CONTROL

867
868 - Build a comprehensive theory adapted to the context and intervention at hand

869 As this pilot example only focused on one mechanism, we cannot refine the comprehensive
870 theory on performance-based financing. We can, however, interpret the refined mechanism
871 concept ‘self-efficacy’ in light of a performance-based financing intervention. This means
872 that we can indicate which program components of a PBF scheme are important to trigger
873 the self-efficacy mechanism which may lead to more intrinsic motivation and better quality
874 of care.

875 What this mechanism concept learns us, is that initial self-efficacy plays an important role
876 in triggering the self-efficacy mechanism. This may explain different outcomes across health
877 workers and may reinforce quality of care inequities between health workers and/or facilities.
878 When initial self-efficacy is absent, the mechanism concept shows that positive experiences,
879 effort and perceived control are crucial in triggering the self-efficacy mechanism.

880 Clearly, when implemented well, PBF can have an important impact on the triggering of the
881 self-efficacy mechanism by creating moments for constructive and positive feedback,
882 increasing the effort of the health workers through financial incentives and improving the
883 work environment by increasing the funds of the health facility and incentivizing the

two solution terms. This means that one prime implicant in this solution term is logically redundant. However, since parsimoniousness is not the ultimate objective here, this is not problematic.

884 facilities. Although this is only an analysis of one mechanism, the relevance of other
885 mechanisms mentioned in figure 1 already becomes clear. For example, the income effect,
886 extrinsic motivation or learning.
887

888 References

- 889 Abner, G. B., Kim, S. Y., & Perry, J. L. (2017). Building Evidence for Public Human Resource
890 Management: Using Middle Range Theory to Link Theory and Data. *Review of Public*
891 *Personnel Administration*, 37(2), 139-159. doi:10.1177/0734371x17697248
- 892 Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human*
893 *Decision Processes*, 50(2), 179-211. doi:[https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- 894 Astbury, B., & Leeuw, F. L. (2010). Unpacking Black Boxes: Mechanisms and Theory Building
895 in Evaluation. *American Journal of Evaluation*, 31(3), 363-381.
896 doi:10.1177/1098214010371972
- 897 Bandura, A. (1994). Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human*
898 *behavior* (Vol. 4, pp. 71-81). New York: Academic Press.
- 899 Befani, B., Ledermann, S., & Sager, F. (2007). Realistic Evaluation and QCA: Conceptual
900 Parallels and an Empirical Application. *Evaluation*, 13(2), 171-192.
901 doi:10.1177/1356389007075222
- 902 Bhaskar, R. (2008 [1975]). *A realist theory of science*. London: Verso books.
- 903 Booth, A., Briscoe, S., & Wright, J. M. (2020). The “realist search”: A systematic scoping
904 review of current practice and reporting. *Research Synthesis Methods*, 11(1), 14-35.
905 doi:<https://doi.org/10.1002/jrsm.1386>
- 906 Booth, A., Harris, J., Croot, E., Springett, J., Campbell, F., & Wilkins, E. (2013). Towards a
907 methodology for cluster searching to provide conceptual and contextual “richness” for
908 systematic reviews of complex interventions: case study (CLUSTER). *BMC Medical*
909 *Research Methodology*, 13(1), 118. doi:10.1186/1471-2288-13-118
- 910 Borghi, J., Singh, N. S., Brown, G., Anselmi, L., & Kristensen, S. (2018). Understanding for
911 whom, why and in what circumstances payment for performance works in low and
912 middle income countries: protocol for a realist review. *BMJ Global Health*, 3(3),
913 e000695. doi:10.1136/bmjgh-2017-000695
- 914 Byrne, D. (2018). Researching complex large-scale nested interventions. In N. Emmel, J.
915 Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist research*
916 (pp. 91-106). London: Sage.
- 917 Cronqvist, L. (2019). QCA Add-In [Version 1.1]. Retrieved from <https://www.qca-addin.net>
- 918 Crossref. (2019). Cross ref Annual Report & Fact File 2018-19. Retrieved from
919 <https://doi.org/10.13003/y8ygwm5>
- 920 Dalkin, S. M., Greenhalgh, J., Jones, D., Cunningham, B., & Lhussier, M. (2015). What’s in a
921 mechanism? Development of a key concept in realist evaluation. *Implementation*
922 *Science : IS*, 10, 49. doi:10.1186/s13012-015-0237-x
- 923 Emmel, N. (2021). Post-disciplinary realism. *International Journal of Social Research*
924 *Methodology*, 24(1), 95-108. doi:10.1080/13645579.2020.1803526
- 925 Fritsche, G. B., Soeters, R., & Meessen, B. (2014). *Performance-Based Financing toolkit*.
926 Washington, D.C.: The World Bank.
- 927 Gerrits, L., & Verweij, S. (2013). Critical Realism as a Meta-Framework for Understanding the
928 Relationships between Complexity and Qualitative Comparative Analysis. *Journal of*
929 *Critical Realism*, 12(2), 166-182. doi:10.1179/rea.12.2.p663527490513071
- 930 Goertz, G. (2006). *Social Science Concepts: A User's Guide*. Princeton: Princeton University
931 Press.
- 932 Goertz, G. (2017). *Multimethod Research, Causal Mechanisms, and Case Studies: An*
933 *Integrated Approach*. Princeton: Princeton University Press.
- 934 Goertz, G., & Starr, H. (Eds.). (2002). *Necessary Conditions: Theory, Methodology, and*
935 *Applications*: Rowman & Littlefield.

- 936 Goicolea, I., Vives-Cases, C., Hurtig, A.-K., Marchal, B., Briones-Vozmediano, E., Otero-
937 García, L., . . . San Sebastian, M. (2015). Mechanisms that Trigger a Good Health-Care
938 Response to Intimate Partner Violence in Spain. Combining Realist Evaluation and
939 Qualitative Comparative Analysis Approaches. *PLoS One*, *10*(8), e0135167.
940 doi:10.1371/journal.pone.0135167
- 941 HRITF. (2015). The Performance-Based Financing (PBF) Conceptual Framework. Retrieved
942 from [https://www.rbfhealth.org/resource/performance-based-financing-conceptual-](https://www.rbfhealth.org/resource/performance-based-financing-conceptual-framework)
943 [framework](https://www.rbfhealth.org/resource/performance-based-financing-conceptual-framework)
- 944 Jagosh, J. (2020). Retroductive theorizing in Pawson and Tilley's applied scientific realism.
945 *Journal of Critical Realism*, *19*(2), 121-130. doi:10.1080/14767430.2020.1723301
- 946 Kalmar, D. A., & Sternberg, R. J. (1988). Theory knitting: An integrative approach to theory
947 development. *Philosophical Psychology*, *1*(2), 153-170.
948 doi:10.1080/09515088808572934
- 949 Leeuw, F. L., & Donaldson, S. I. (2015). Theory in evaluation: Reducing confusion and
950 encouraging debate. *Evaluation*, *21*(4), 467-480. doi:10.1177/1356389015607712
- 951 Lohmann, J., Houllfort, N., & De Allegri, M. (2016). Crowding out or no crowding out? A Self-
952 Determination Theory approach to health worker motivation in performance-based
953 financing. *Social Science and Medicine*, *169*, 1-8. doi:10.1016/j.socscimed.2016.09.006
- 954 Marchal, B., Kegels, G., & Van Belle, S. (2018). Theory and realist methods. In N. Emmel, J.
955 Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist research*
956 (pp. 79-89). London: Sage.
- 957 Merton, R. K. (1968). *Social theory and social structure*. New York: Free Press.
- 958 Nimpagaritse, M., Korachais, C., Roberfroid, D., Kolsteren, P., Zine Eddine El Idrissi, M. D., &
959 Meessen, B. (2016). Measuring and understanding the effects of a performance based
960 financing scheme applied to nutrition services in Burundi—a mixed method impact
961 evaluation design. *International journal for equity in health*, *15*(1), 93.
962 doi:10.1186/s12939-016-0382-0
- 963 Pawson, R. (2006a). Digging for Nuggets: How 'Bad' Research Can Yield 'Good' Evidence.
964 *International Journal of Social Research Methodology*, *9*(2), 127-142.
965 doi:10.1080/13645570600595314
- 966 Pawson, R. (2006b). *Evidence-Based Policy: A Realist Perspective*. London: SAGE
967 Publications.
- 968 Pawson, R. (2013). *The science of evaluation: A realist manifesto*. London: Sage publications.
- 969 Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2004). Realist synthesis: an
970 introduction. *ESRC Research Methods Programme Working Paper*, *2/2004*.
- 971 Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London: Sage.
- 972 Perry, J. L. (2010). A Strategic Agenda for Public Human Resource Management Research.
973 *Review of Public Personnel Administration*, *30*(1), 20-43.
974 doi:10.1177/0734371x09351821
- 975 Ragin, C. C. (2009). *Redesigning social inquiry: Fuzzy sets and beyond*. Chicago: University of
976 Chicago Press.
- 977 Ragin, C. C. (2014 [1987]). *The Comparative Method: Moving Beyond Qualitative and*
978 *Quantitative Strategies*. Berkeley, CA: University of California Press.
- 979 Renmans, D., Holvoet, N., Criel, B., & Meessen, B. (2017). Performance-Based Financing: the
980 same is different. *Health Policy and Planning*, *32*(6), 860-868.
981 doi:10.1093/heapol/czx030
- 982 Ridde, V., Turcotte-Tremblay, A. M., Souares, A., Lohmann, J., Zombré, D., Koulidiati, J. L., .
983 . . De Allegri, M. (2014). Protocol for the process evaluation of interventions combining
984 performance-based financing with health equity in Burkina Faso. *Implementation*
985 *Science*, *9*, 149. doi:10.1186/s13012-014-0149-1
- 986 Rutten, R. (2021). Uncertainty, possibility and causal power in QCA. *Sociological Methods &*
987 *Research*. doi:Not yet available
- 988 Sager, F., & Andereggen, C. (2012). Dealing With Complex Causality in Realist Synthesis:The
989 Promise of Qualitative Comparative Analysis. *American Journal of Evaluation*, *33*(1),
990 60-78. doi:10.1177/1098214011411574

991 Sayer, A. (1992). *Method in Social Science: A Realist Approach*. London, New York:
992 Routledge.

993 Sayer, A. (2000). *Realism and Social Science*. London: SAGE Publications.

994 Schneider, C. Q., & Wagemann, C. (2012). *Set-Theoretic Methods for the Social Sciences: A*
995 *Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.

996 Tomoaia-Cotisel, A., Kim, H., Allen, S. D., & Blanchet, K. (2017). Causal loop diagrams: a tool
997 for visualizing emergent system behaviour. In D. de Savigny, K. Blanchet, & T. Adam
998 (Eds.), *Applied systems thinking for health systems research: A methodological*
999 *handbook* (pp. 97 - 114). London: Open University Press.

1000 Westhorp, G. (2018). Understanding mechanisms in realist evaluation and research. In N.
1001 Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist*
1002 *research* (pp. 41-57). London: Sage Publications.

1003 Williams, M. (2018). Making up mechanisms in realist research. In N. Emmel, J. Greenhalgh,
1004 A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist research* (pp. 25-40).
1005 London: Sage.

1006 Wong, G. (2018). Data gathering in realist reviews: Looking for needles in haystacks. In N.
1007 Emmel, J. Greenhalgh, A. Manzano, M. Monaghan, & S. Dalkin (Eds.), *Doing realist*
1008 *research* (pp. 131-145). London: Sage.

1009

1010