


RESEARCH

Open Access



Policymakers' experience of a capacity-building intervention designed to increase their use of research: a realist process evaluation

Abby Haynes^{1,2*} , Sue Brennan³, Sally Redman¹, Anna Williamson¹, Steve R. Makkar¹, Gisselle Gallego⁴ and Phyllis Butow⁵

Abstract

Background: An intervention's success depends on how participants interact with it in local settings. Process evaluation examines these interactions, indicating why an intervention was or was not effective, and how it (and similar interventions) can be improved for better contextual fit. This is particularly important for innovative trials like Supporting Policy In health with Research: an Intervention Trial (SPIRIT), where causal mechanisms are poorly understood. SPIRIT was testing a multi-component intervention designed to increase the capacity of health policymakers to use research.

Methods: Our mixed-methods process evaluation sought to explain variation in observed process effects across the six agencies that participated in SPIRIT. Data collection included observations of intervention workshops (n = 59), purposively sampled interviews (n = 76) and participant feedback forms (n = 553). Using a realist approach, data was coded for context-mechanism-process effect configurations (retroductive analysis) by two authors.

Results: Intervention workshops were very well received. There was greater variation of views regarding other aspects of SPIRIT such as data collection, communication and the intervention's overall value. We identified nine inter-related mechanisms that were crucial for engaging participants in these policy settings: (1) Accepting the premise (agreeing with the study's assumptions); (2) Self-determination (participative choice); (3) The Value Proposition (seeing potential gain); (4) 'Getting good stuff' (identifying useful ideas, resources or connections); (5) Self-efficacy (believing 'we can do this!'); (6) Respect (feeling that SPIRIT understands and values one's work); (7) Confidence (believing in the study's integrity and validity); (8) Persuasive leadership (authentic and compelling advocacy from leaders); and (9) Strategic insider facilitation (local translation and mediation). These findings were used to develop tentative explanatory propositions and to revise the programme theory.

Conclusion: This paper describes how SPIRIT functioned in six policy agencies, including why strategies that worked well in one site were less effective in others. Findings indicate a complex interaction between participants' perception of the intervention, shifting contextual factors, and the form that the intervention took in each site. Our propositions provide transferable lessons about contextualised areas of strength and weakness that may be useful in the development and implementation of similar studies.

Keywords: Participant perspectives, Research utilisation, Process evaluation, Realist evaluation, Health policy

* Correspondence: abby.haynes@saxinstitute.org.au

¹Sax Institute, 235 Jones Street, Ultimo, NSW 2007, Australia

²Sydney School of Public Health, Edward Ford Building (A27), University of Sydney, Camperdown, NSW 2006, Australia

Full list of author information is available at the end of the article

Background

This paper presents a realist analysis of how a novel, multi-component intervention trial designed to increase research use capacity, known as the Supporting Policy In health with Research: an Intervention Trial (SPIRIT), functioned in six health policy agencies. Data from a mixed-methods process evaluation is used to unpack the processes of engagement and participation that were hypothesised to mediate the intervention's success. These intermediate impacts are conceptualised as process effects (see Box 1 for definitions). We do this by describing what was delivered in the intervention and what process effects were observed, then identify explanatory 'Context + Mechanism → Process effect' configurations that show how the intervention, and the trial more broadly, was perceived by participants, why this varied across the participating organisations, and how these perceptions affected receptivity to the intervention's ideas and resources. A realist approach is used because it supports rigorous comparative analysis of how those targeted by an intervention make sense of what it offers, and how this is shaped by context [1–3].

Box 1 Definitions of key concepts used in this paper

Context	In realist terms, context is any system, structure or condition that affects outcomes, including individuals' attributes and social interactions [3]
Mechanism	Mechanisms are what makes an intervention work: "They are not the observable machinery of program activities, but the response that interaction with a program activity or resource triggers (or does not trigger) in the reasoning and behaviour of participants" [70]
Process effects	These are proximal impacts that influence intervention outcomes or are of evaluative interest for other reasons (e.g. they help explain unexpected variation in implementation); others use the term 'formative outcomes' [84]; Desired process effects are those the investigators consider to be prerequisites for a successful intervention
Programme theory	This is, "An explicit theory or model of how an intervention contributes to a set of specific outcomes through a series of intermediate results" [85]; programme theory should be plausible, useful and consistent with the evidence
Proposition	Propositions are generalised theoretical statements grounded in the data [86]; in realist evaluation, they link and condense information about contexts, mechanisms and outcomes; propositions are refined through empirical testing but remain fallible [87]
Realist process evaluation	Process evaluation helps explain how an intervention had its effects [7]; realist process evaluation applies realist principles to this process and investigates causal patterns (known as demi-regularities) to show how intervention strategies may be operating under what conditions to generate process effects for which groups [3]
Retroduction	This is a form of analysis that "involves constant shuttling between theory and empirical data, using both inductive and deductive reasoning" [88]

Understanding interventions

Interventions – planned activities to change individual, group and/or organisational behaviour – are not passively received, but are actively shaped by the people who participate in them and the circumstances in which they are delivered [4–6]. Understanding the ways in which participants interact with and perceive an intervention is vital for determining how and why it was, or was not, effective [7]. This requires moving beyond measures of participant satisfaction – sometimes derided as "happy face evaluation" [8] – towards methods which delve into "the complexity, flux and contextual variation that inevitably occurs in real life situations" [9].

Many organisational capacity-building interventions fail because they do not take sufficient account of participants' workplaces [10]. Successful interventions introduce strategies (ideas, activities and resources) that are contextually apt [7, 11] and which are therefore able to produce desired interactions [3]. For example, in organisational interventions, participants' perceptions and interactions are affected by factors such as the organisation's culture [12], its history of change [13, 14], staff heterogeneity [15] and trust in management [13].

Information about how implementation interacts with people and place over the course of an intervention is frequently overlooked [16]; yet, it is necessary for making informed assessments about the worth, adaptability and transferability of strategies designed to bring about individual or organisational change [9]. In multi-component interventions it is often impossible to disentangle which components were more or less effective, or what variations in combination might maximise effectiveness [17]. These interventions frequently trigger unanticipated causal processes and have unpredictable impacts that standardised measures are unlikely to capture [18]. This may be especially important for interventions where participants have involvement in the tailoring and/or delivery of an intervention, since their attitudes towards its content, form and goals are likely to have profound impacts on what is delivered and how it is received [19, 20]. Indeed, there is an established link between outcomes and the ways that participants gauge the quality of their involvement in tailoring the scope, content and process of flexible interventions [4].

Context-sensitive design, implementation and evaluation are particularly pressing for interventions that attempt to increase the use of research in policy processes. Policymaking is "a contested arena of negotiation.... messy, complex, and serendipitous" [21], (where research, and researchers [22]), are used strategically [23, 24]. Macro-level political and institutional factors influence how policy-makers and policy organisations engage with and make use of research [23], and will therefore mediate their relationships with research utilisation interventions. Given that the use of research is cultural and rhetorical as well as technical

[25], where an intervention promotes greater use of research, or claims to be evidence based, participants may actively critique that premise [26, 27]. Thus, determining if and how such an intervention is compatible with participants' beliefs and practice norms is critical.

Despite these arguments, many interventions are reported (and, by implication, conducted) with minimal consideration of the interactions between the intervention activities, the people who took part, and the circumstances that mediated this relationship [9, 28]. As Clark et al. note, "Little research has explored individuals' experiences of programmes or examined how programme dimensions lead to changes in behaviour. ...individuals' meanings, experiences and reactions to the programme and the effects of their wider context are simultaneously disregarded" [29]. Realist process evaluation is well equipped to redress these oversights [1, 3].

The study being evaluated: SPIRIT

SPIRIT was a stepped wedge cluster randomised trial that tested the effects of a novel intervention designed to increase the capacity of health policy agencies to use research. Six organisations in Sydney, Australia, participated. Five were state government agencies and one was a national organisation funded by the federal government. An agency was eligible to participate if (1) a significant proportion of its work was in health policy or programme development, and (2) there were at least 20 staff involved in health policy, programme development or evaluation. A sampling frame was drawn from Government websites that listed all New South Wales and Australian government health policy and programme agencies located in Sydney. Members of the investigator team reduced this list to 16 potentially eligible agencies

and ranked as highest those with the greatest focus on health and the largest numbers of relevant staff. The top six agencies were invited to take part, and all agreed [30]. Each agency's Chief Executive Officer (CEO) signed an organisational-level agreement to participate in SPIRIT and nominated a liaison person: an internal member of staff who would be responsible for coordinating SPIRIT in their setting for the duration of the trial. There were six rounds of outcome data collection using three evaluation tools. These are described in detail elsewhere [30–35].

The intervention aimed to increase agency capacity to use research in relation to three goals, namely (1) the organisation and staff value research more; (2) more tools and systems are in place to support research engagement actions and the use of research; and (3) staff have greater knowledge and skill in research engagement actions and the use of research. SPIRIT's design was informed by an action framework [36] and underpinning change principles that reflected composite theory from psychology, organisational science, adult learning and the research utilisation literature [30]. The intervention comprised multiple components hinging on interactive workshops such as research skills seminars, exchange forums with researchers, and a leadership programme targeting senior managers. Other activities included the provision of tools and resources (such as an online research portal); practice using systems for commissioning research reviews, analyses or evaluations; and CEO espousal of research-informed policymaking (Fig. 1). Agencies could choose options within and tailor many of the components to address local priorities. Each agency was asked to identify two lists of potential participants, namely (1) all staff involved in policy or programme development, implementation or evaluation who would be

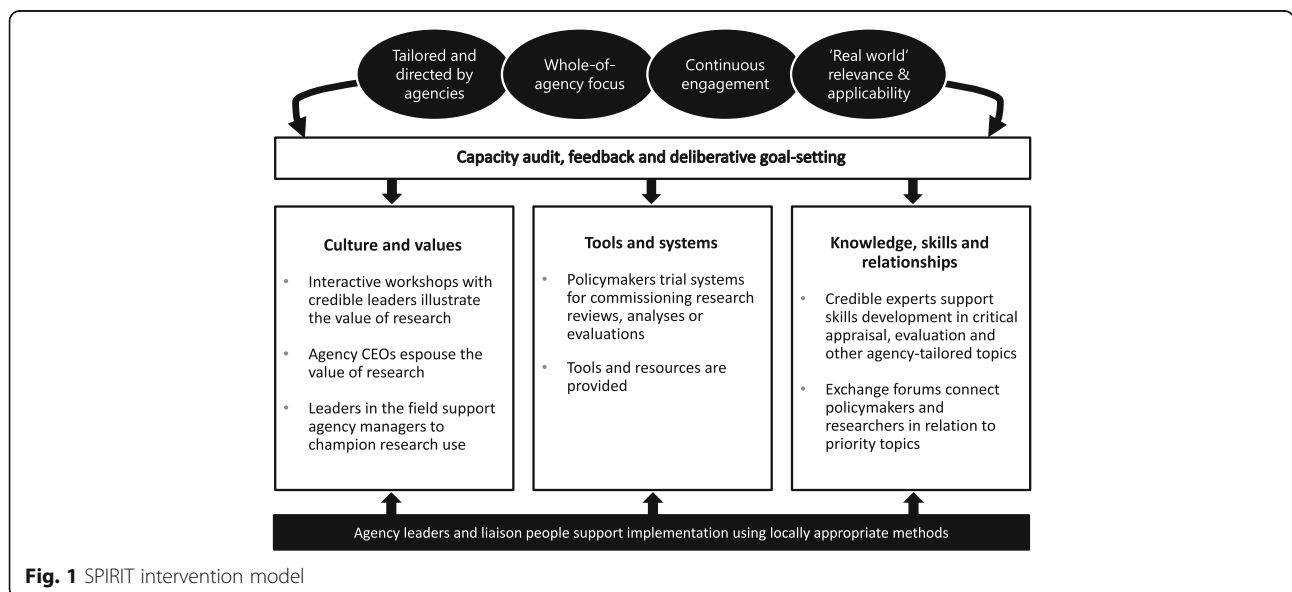


Fig. 1 SPIRIT intervention model

invited to take part in intervention activities and data collection and (2) managers who would take part in the leadership programme and promote SPIRIT.

An onsite introductory information session preceded the intervention and data collection in each site. The round of data collection that took place immediately before the intervention functioned as an audit and was followed by a feedback forum in which the lead investigator facilitated a deliberative dialogue with leaders about their agency's findings. Intervention goals targeting research engagement and use were identified during this process. Agency leaders considered how they would like to use SPIRIT's options to address these goals and, if applicable, any additional (non-SPIRIT) strategies for reaching their goals.

External research and policy experts were contracted to deliver workshops. They were briefed on SPIRIT's 'change principles' and their workshop's objectives. The content of the tailored workshops was negotiated with the agency's liaison person, with input from presenters. Members of the SPIRIT research team coordinated the development and delivery of workshops and other intervention activities. Each site had a dedicated knowledge broker from the SPIRIT team who acted as the onsite 'face' of SPIRIT, negotiated tailoring and attended all intervention activities.

An in-depth, mixed methods process evaluation informed by realist thinking was conducted in parallel with the intervention. This paper is based on that data.

The role of process evaluation

Process evaluation investigates an intervention's implementation, change mechanisms and contextual interactions in order to explain (insofar as this is possible) how and why the intervention functioned as it did in each intervention site [18]. Process evaluation does not determine whether study outcomes are achieved, but it can identify process effects, namely proximal impacts of an intervention that make achieving outcomes more or less likely [37].

Aims

Using a realist evaluation approach [1, 3, 38, 39], we aimed to generate transferable learning in relation to the questions, (1) To what extent did SPIRIT achieve the desired process effects in each agency? and (2) How were these process effects generated? i.e. What mechanisms seem to account best for the patterns of engagement and participation observed across all agencies?

Methods

Realist evaluation

The SPIRIT process evaluation comprised a fidelity assessment and a theory-driven exploration of the interaction between the intervention, participants and the implementation circumstances, with the expectation that this would probably take a different form in each of the

six agencies [40]. Theory-driven evaluation seeks to uncover causal pathways [41] and is well suited for understanding how multicomponent interventions function in complex real-world settings [42]. In this study, we adopt a particular theory-driven approach – a realist evaluation [43] – following the methods associated with Pawson [1], Pawson and Tilley [3], and others in the RAMESES II project [39]. Realist evaluation focuses on an intervention's underlying theory as its unit of analysis [1, 3], with the aim of determining “*what works, for whom, in what circumstances, and how*” [3, 44]. Realists posit that interventions introduce ideas and opportunities that generate effects in conjunction with participants' reasoning and resources. Thus, the interaction between intervention activities and the contexts of each intervention site will determine what (if any) mechanisms are activated and what outcomes (intended and unintended) are generated [45, 46].

We used a realist approach because it maximises the transferability of findings and operational learning from one setting to another (an enduring concern in intervention evaluation [47]), while also recognising complexity and the need to look beyond one-size-fits-all ways of responding to problems [1, 3, 48, 49]. Realist evaluation has been used effectively in studies of policy processes [50], implementation research [51], knowledge exchange [52] and evaluations of flexible intervention trials [19, 29], making it especially suitable for addressing the methodological challenges presented by a multi-component, novel and theoretically eclectic trial like SPIRIT (outlined in detail elsewhere [53]).

Importantly, analyses arising from realist evaluations are tentative, claiming only to be an informed hypothesis of “*how something might be*” [54] rather than a definitive version of reality. These hypotheses accrue plausibility when tested in further studies, but remain open to revision or rejection if alternative theories are more convincing [45]. In our study, data collection, management and analysis were concurrent; thus, we were continually testing and revising hypotheses within and across the six intervention sites over the 30-month study, but our findings are embryonic in realist terms.

Initial programme theory

Realist evaluation develops, tests and refines programme theory. SPIRIT was informed by a mixture of formal theory and experiential knowledge [30], and had both a well-articulated action framework [36] and clear principles about what should be provided [53], but did not offer hypotheses about the mechanisms that would generate increased capacity to use research. Based on existing trial materials and discussions with the designers, we articulated the overarching programme theory to make the intended causal pathway more explicit so that we could critique the assumptions underpinning the intervention design [1, 3, 55]. This was refined and agreed through further consultation:

SPIRIT will engage and motivate agency leaders to 'own' the intervention using audit feedback, deliberative goal-setting and program tailoring. This agency-driven approach will generate a priority-focused program that offers locally relevant practice support and accommodates differences in agencies' values, goals, resources and remits. The program will comprise a suite of andragogical activities, tools, and connection across the research-policy divide that provide resources and build knowledge, skills and relationships. It will be supported via modelling and opinion leadership by agency leaders and dynamic external experts. CEOs will promote SPIRIT in their agencies and liaison people will facilitate the tailoring and implementation. These strategies will act synergistically to stimulate and resource participants at different organisational levels, leading to changes in values, practice behaviours and agency processes. This will facilitate increased use of research in policy processes.

This pathway informed the data collection, providing pointers about what to look for, but was used flexibly (rather than as a rigid investigative framework) as befits an exploratory study. We also looked for unintended effects, and considered alternative causal pathways that might better explain observed effects. The data offered the opportunity to develop a much richer understanding of the social processes and interactions than had previously been possible.

Process effects

The programme theory was used to identify desired process effects via discussion with the study designers. We then explored how these process effects were achieved in each setting for the range of targeted participants, or why they were not. Our conceptual framework for this work was informed by the implementation science literature that focuses on social processes and interaction in interventions (e.g. [6, 26, 56–60]).

Data collection

Causation, and the mechanisms that generate it, are seldom observable [3]. Therefore, in realist evaluation, data is triangulated to identify the interactive patterns that can most plausibly explain how the intervention led to the observed outcomes [61]. Quantitative data is helpful for identifying outcomes [1], while qualitative methods are usually necessary “to discover actors' reasoning and circumstances in specific contexts” [62]. We used the following methods to capture information:

- Semi-structured interviews with 5–9 participants from each agency early in the intervention period (n = 33) and post-intervention (n = 43). Interviewees were purposively selected for maximum variation

in work roles, attitudes to research and experiences of SPIRIT in order to explore the breadth of dimensions expected to influence interactions with the intervention [7]. Open-ended questions and prompts explored interviewees' work practices and contexts, and their experiences and perceptions of SPIRIT, including their explanations for any change. The interview questions are available elsewhere [40]. This combination of context-, causal- and impact-focused questions across diverse participants was used to refine theory about what was working (or not), for whom and in what circumstances.

- Observations of intervention workshops (n = 59), and informal opportunistic conversations with participants before and after workshops. Workshops were audio recorded and field notes were written immediately afterwards. A checklist was used for fidelity coding through which we monitored the extent to which 'essential elements' of the intervention were delivered (detailed elsewhere [59]).
- Anonymous participant feedback forms (n = 553). These comprised Yes/No ratings on six statements: (1) The workshop was interesting, (2) The workshop was relevant to my work, (3) The workshop was realistic about the challenges and constraints of our work, (4) The presenter had appropriate knowledge and skills, (5) It is likely that I will use information from this workshop in my work, (6) It is likely that SPIRIT will benefit my agency (Additional file 1). Some workshops had additional items, e.g. the forms for audit feedback forums included items about the clarity of the data and participants' confidence that SPIRIT would be adequately tailored for their agency. All forms contained three open-ended questions: (1) 'What worked well?', (2) 'What could be improved?' and (3) 'Any other comments?' Forms were distributed prior to intervention workshops and completed immediately afterwards.
- Formal and informal interviews with the people implementing SPIRIT and the commissioned presenters.
- Limited access to information from the interviews conducted as part of SPIRIT's outcome evaluation. These interviews focused on (1) organisational support for research use (n = 6), and (2) the role of research in the development of a recent policy or programme (n = 24). We reviewed transcripts from the first round of interviews (prior to the intervention), but thereafter were blinded to this data so that it would not influence the ongoing process evaluation analysis.

Data management and analysis

Qualitative data

Data was initially analysed for the whole process evaluation. Interview data was managed using framework

analysis [63] within NVivo v.10 [64] and used to develop descriptive case studies [65] in combination with data from the fidelity assessment, running memos for each agency, interviewee memos, the thematically coded data from field notes and the open-ended questions in feedback forms. These case studies described (1) each agency's context, change trajectory, workforce and practice norms, (2) their research use practices and culture, (3) how SPIRIT was implemented in each setting, and (4) the interactions between (1), (2) and (3). Framework categories and the structure of the case studies were iteratively developed from *a priori* concerns (such as the constructs the intervention was targeting and the hypothesised causal pathway), and from themes identified using inductive analysis [66, 67]. The method of constant comparison [68] was used to query and refine the initial programme theory and other emergent hypotheses throughout the trial. This work is described in more detail elsewhere [40].

Quantitative data

For each agency, we calculated the number and percentage of feedback forms responding 'Yes' to each of the six statements outlined earlier. In calculating these frequencies, the four different types of workshops (symposia, research exchanges, leaders' forums and audit feedback forums) were aggregated.

Realist analysis

Using the data described above, we sought to explore the hypothesised pathway identified in the initial programme theory and to identify any other pathways leading to the interventions' observed process effects, plus other impacts reported by participants or members of the implementation team [42].

We employed a retroductive analytical approach that attempts to explain phenomena by theorising about what mechanisms are capable of producing them [69]. This involves studying events "*with respect to what may have, must have, or could have caused them. In short it means asking why events have happened in the way they did*" [51]. In accordance with realist evaluation principles, we focused on the interaction of SPIRIT with features of each agency's context that appeared most likely to have influenced process effects [42, 70]. We developed explanatory configurations of the patterns we saw in the data. In realist evaluation, these are typically called Context + Mechanism → Outcome configurations [1, 3], but because the 'outcomes' of interest in process evaluation are process effects rather than study outcomes, we have called them Context + Mechanism → Process effect configurations herein. Propositions were then developed to summarise each configuration. This work depended on using each type of data to query, explain and balance the other to reach as comprehensive as possible

accounts of what happened and why [71, 72]. Original data sources were revisited as required.

These process effects were identified prior to the development of Context + Mechanism → Process effect configurations and were used as a starting point in much of the analysis – although realist evaluation depicts outcomes (or, in our case, process effects) as the final step in the sequence, the analysis tends to start by identifying effects, then working backwards to investigate the conditions (context and mechanisms) that caused them [73]. We traced connections to and from observed process effects asking 'What caused this?', 'Why didn't this unfold as anticipated?' and 'What best explains these different responses between agencies?' Analysis involved looking for data that might indicate the absence or weak functioning of mechanisms as well as the presence of a mechanism. This was aided by Dalkin et al.'s [46] assertion that mechanisms may vary in intensity rather than simply being present or absent.

AH, who led the process evaluation, reviewed and coded all data sources. SB, who contributed to the process evaluation design and analysis throughout the trial, independently reviewed a proportion of interview transcripts and cross-agency fieldwork memos. Their preliminary Context + Mechanism → Process effect configurations overlapped extensively and were workshopped with further reference to the wider data set to develop agreed configurations. Further discussion with our co-authors resolved differences and refined the final findings.

This analysis relied on abductive reasoning [74], which is an iterative cycling between data and likely explanations that incorporates inductive and deductive processes. We looked for evidence of factual causal mechanisms, and for evidence that supported, discounted or nuanced current causal hypotheses both in real time (as the intervention unfolded) and retrospectively (reviewing data already collected). Throughout this process, we sought to identify where our evolving Context + Mechanism → Process effect configurations aligned with existing theory; we revisited the theories used to inform the development of SPIRIT, asking to what extent did these theories support the patterns we were observing in the data, and also considered other theories that might better explain our findings. See Additional file 2 for an overview.

Results

In this section, we describe the implementation of the SPIRIT intervention, outline the observed process effects, and then attempt to explain how these effects were generated using Context + Mechanism → Process effect configurations. Finally, we present the revised programme theory.

Implementation

As Additional file 3 shows, some aspects of SPIRIT were delivered with a high degree of implementation fidelity;

Table 1 Overview of SPIRIT's process effects and data sources

Desired process effects for the trial	Observed process effects	Supporting data sources
1. Leaders espouse SPIRIT and its goals	All CEOs disseminated initial information about their agency's participation in SPIRIT, but only four had a continuing visible role in supporting the intervention, e.g. sending updates and attending workshops; some executive members participated in each site, but to very different extents ranging from a half hour 'drop in' to repeated and enthusiastic participation; many managers talked about SPIRIT in team meetings and encouraged their staff to attend	Interviews at two time points (early-intervention 'context' and post-intervention 'perceptions and impact'), ad hoc conversations with participants
2. Liaison people facilitate the intervention effectively	The use of a liaison person was very effective in the sites where the liaison person was enthusiastic about SPIRIT; four of the six worked hard to promote, tailor and administer the intervention, harnessing insider knowledge and using creative strategies, whereas the other two did not tailor or promote the intervention as thoroughly and expressed negative views to colleagues about SPIRIT	Observations of workshops, interviews and conversations as above, feedback from the SPIRIT team about their communications with liaison people
3. Targeted policymakers participate in, and are receptive to, intervention activities	Participation levels were good in that they met the SPIRIT team's expectations for each site; each agency targeted different groups for different components so proportions and types of participants varied, but liaison people were satisfied with attendance and were occasionally surprised by very high numbers; attendance at workshops averaged between 11 and 20 participants per workshop, with between 102 and 158 total occasions of attendance across the six sites; there was full participation in other activities (e.g. trialling the commissioned research services); receptivity varied tremendously within, but especially between, agencies: see next section for more details, including possible reasons	Quantitative fidelity data from observations (using check lists and sign-in sheets), observations, interviews and conversations as above
4. Participants actively contribute to the content of those activities	Where there was opportunity, participants contributed greatly to workshop content via questions, discussion and case examples; interactivity was limited on some occasions in all agencies, usually because the presenter provided few opportunities; in larger groups, more senior staff tended to dominate, but other participants said this was still useful. Some liaison people helped craft workshop content and provided agency-based case examples; one agency co-presented a workshop; the agency staff nominated to test the research commissioning service were actively involved	Observations of workshops, including descriptive accounts of interactions and dynamics
5. Participants identify potentially useful ideas, techniques and/or resources	94% of those who completed a feedback form said they found workshops to be both relevant to their work and realistic about policy challenges and constraints; many interviewees identified specific benefits from SPIRIT, including improved awareness of useful researchers and research resources, understanding of the evidence relating to a policy problem and access to existing agency resources	Participant feedback forms, observations of workshops, interviews and ad hoc conversations with participants and liaison people
6. Participants use, or plan to use, these ideas, techniques and/or resources	Workshops facilitated less discussion than intended about how learning might be applied, but 95% of participants who completed a feedback form agreed, " <i>It is likely that I will use information from this workshop in my work</i> "; some interviewees said they planned to use ideas or resources, and a few had done so, especially newer staff; three liaison people had managerial-approved plans underway for research-focused education and/or systems improvement, e.g. mandated consideration of research in policy proposals; two agencies had plans to use their commissioned research products	
Desired process effects for the evaluation	Observed process effects	Supporting data sources
7. Liaison people facilitate data collection effectively	All liaison people facilitated data collection sufficiently, although it was occasionally delayed and required prompting; where liaison people championed SPIRIT they used additional strategies to encourage participation in data collection, in one agency this achieved a 100% response rate	Outcome measures completion figures, interviews with participants and liaison people, feedback from SPIRIT team
8. Targeted participants take part in data collection	In all agencies, there was full participation in the two interview-based measures, but more variable responses to the anonymous online survey; response rates dipped in the second measurement point, but stabilised after the survey was shortened; overall, the online survey response rate was 56% and there was a mean 74% response rate for process evaluation feedback forms; only three-quarters of invitees took part in a process evaluation interview	Outcome measures completion figures, interviews with participants and liaison people
9. The benefits of the intervention are judged to outweigh the burdens of the trial	Interviewees differed considerably in their assessments of the intervention, but where they felt it had benefits these were deemed to outweigh the trial's burdens, this included those liaison people who championed SPIRIT from the start; workshops with high profile and dynamic 'service-orientated' presenters were especially valued; nearly 98% of all feedback form respondents agreed with the statement, " <i>It is likely that SPIRIT will benefit my agency</i> "	Early-intervention and post-intervention interviews, ad hoc conversations with participants and liaison people, feedback form data

indeed, every agency received audit feedback and the intended number of components on the topics they requested. Intra-organisational processes that were outside the control of the implementation team had greater variation. The promotion of SPIRIT and much of its administration depended on the attitudes and behaviours of liaison people and each organisation’s leaders, and to a lesser extent, the expert presenters commissioned for each workshop. This resulted in some loss of SPIRIT’s theoretical fidelity, i.e. the extent to which the intervention delivered its ‘essential elements’ (these are discussed in more detail elsewhere [53]). For example, the essential elements stipulated that workshops should be non-didactic and therefore the presenters should encourage participants to contribute as much as possible. Many workshops were highly interactive, such as the deliberative audit feedback forums, but others were not. This was because (1) the expert presenters sometimes overrode their briefing to facilitate discussion; (2) liaison people occasionally tried to maximise value by cramming content into workshops, which limited opportunities for participation; and (3) unexpectedly, the agencies seldom took up offers to co-design and co-present workshops.

In some sites, SPIRIT’s reach was constrained more than anticipated. Agency 6, for example, chose to focus some components of the intervention on one group of staff and limited participation accordingly. In Agency 3, managers attempted to minimise the onerousness of data collection by excluding some eligible staff from invitations to complete surveys. Agencies also defined their leadership groups quite differently, resulting in wide variation in the

numbers and organisational roles of participants in the leaders’ programme.

Process effects

Table 1 describes SPIRIT’s process effects, i.e. the actions, behaviours and responses hypothesised to be necessary for SPIRIT to generate the capacity-related outcomes measured in the trial. Column 1 lists the process effects both for the intervention and the trial evaluation; we include the latter because of their impact on the quality of the evaluation and the way that SPIRIT as a whole was perceived. Column 2 presents a summary of our observations about the extent to which these process effects occurred. Column 3 shows the data sources for our observations.

How were these process effects generated?

We identified nine primary causal mechanisms (Fig. 2). The Context + Mechanism → Process effect configurations for each mechanism are presented in the following section. Each of the configurations begins with an overview of the context pertaining to that mechanism, a description of how we believe the mechanism functioned, how it generated process effects and how process effects differed between participating agencies. A proposition that summarises the hypothesised casual pathway precedes each configuration.

Cross-references to other mechanisms are in shorthand so that mechanism 1 reads as M1, etc. Similarly, agency numbers are shortened so that Agency 1 is shown as A1, and so on. Inevitably, this is a highly truncated presentation of our findings. For those who seek more

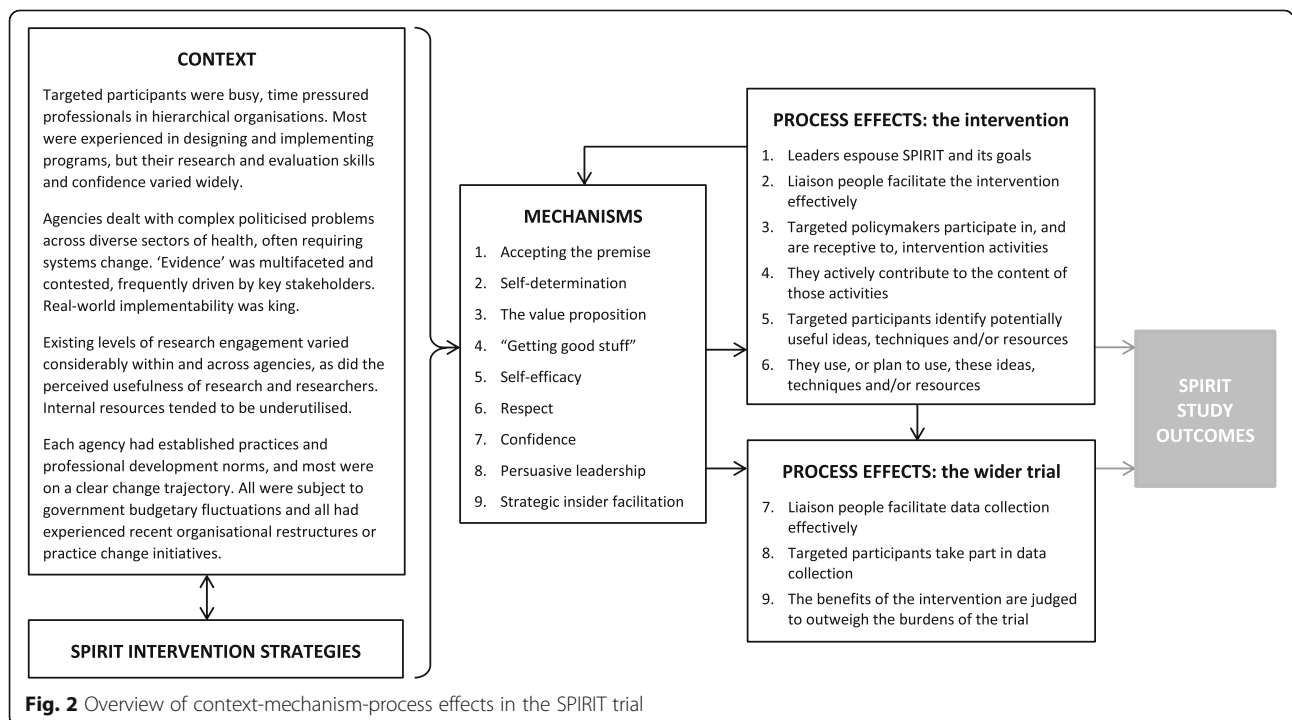


Fig. 2 Overview of context-mechanism-process effects in the SPIRIT trial

detail, a narrative description of the data that informed our identification of each mechanism can be found in Additional file 2. This additional file provides an ‘evidence link’ between the data and the findings that follow.

Mechanism 1

Accepting the premise (Table 2)

Mechanism 2

Self-determination (Table 3)

Mechanism 3

The value proposition (Table 4)

Mechanism 4

“Getting good stuff” (Table 5)

Mechanism 5

Self-efficacy (Table 6)

Mechanism 6

Respect (Table 7)

Mechanism 7

Confidence (Table 8)

Mechanism 8

Persuasive leadership (Table 9)

Mechanism 9

Strategic insider facilitation (Table 10)

Table 2 Mechanism 1 - Accepting the premise

Proposition: Where participants regard the intervention’s form, goals and assumptions as compatible with their agency’s remit, values, practice norms, trajectory of change and current priorities—and providing there is perceived room for improvement—they accept the premise of the study and are receptive to what it offers.

Context	How did mechanism 1 function?	How did mechanism 1 generate process effects?	Agency comparisons
Each agency had existing goals, values, resources, practices and change trajectories. They viewed their capacity to use research, and the importance of increasing this capacity, quite differently. There were diverse norms about what evidence is and how it should be developed, which were affected by their primary stakeholder groups. In the wider environment there was increasing emphasis on <i>action-based research</i> (partnering with practitioners to produce research fit for immediate decision-making). ^a	Potential participants accepted or rejected the premise of SPIRIT based on: 1. The <i>compatibility of SPIRIT’s goals & assumptions</i> with local conceptualisations of evidence and its role in policymaking, including how research related to the agency’s remit, values and practice norms; stakeholder relationships; and change trajectory. 2. The <i>compatibility of SPIRIT’s form</i> – whether its design was congruent with local conceptualisations of ‘good’ or ‘appropriate’ intervention/trial models. 3. <i>Relative advantage</i> : if participants believed they or their agency would benefit from increased use of research, i.e. they saw a need for SPIRIT. 4. <i>Relative priority</i> : they saw this need as immediate, i.e. there was some urgency. <i>Accepting the premise</i> functioned on a continuum. Many policymakers expressed uncertainty rather than a firm view, and modified their view (usually becoming more accepting) while the intervention was underway. Participants’ conception of SPIRIT’s premise did not always align with the designers’ conception.	When they accepted the premise of SPIRIT leaders were more inclined to espouse the intervention and its goals. Liaison people facilitated the intervention more effectively, and targeted policymakers were more enthusiastic about participation and receptive to content. Where policymakers dismissed SPIRIT’s premise they said they were unenthusiastic about participation (yet many did participate) and had low expectations of content. Liaison people who rejected the premise admitted they did not champion SPIRIT or facilitate the intervention as effectively as they could have, but they managed data collection satisfactorily. Leaders who had reservations about the premise tended to express their doubts to staff, but also encouraged staff to participate in specific intervention activities.	Most potential participants in A1, A5 and A6 saw SPIRIT as addressing “ <i>a real need</i> ” and were open to what it had to offer. A3 staff supported the premise but many felt it did not apply to them as they had “ <i>no room for improvement</i> ”. A few interviewees in all agencies dismissed the premise of SPIRIT, but especially in A2 and A4. However, many of these participated in at least one workshop, either because it was expected or because the potential merits of individual workshops (M3) overcame reservations about the intervention/trial as a whole.

^a Edwards M, Evans M: *Getting evidence into policy-making: parliamentary triangle seminar report*. In ANZSIG Insights: ANZOG Institute for Governance, University of Canberra; 2011.

Table 3 Mechanism 2 – Self-determination

Proposition: Where participants have scope and support to shape the form and function of the intervention, and where they make efforts to do so, they will (a) positively enhance the relevance and applicability of content and (b) invest in the intervention. Involvement in shaping even limited aspects of the intervention can deliver benefits. Participants also need to have real choices about participation. Self-determination is linked to M4, M5 and M6.

Context	How did mechanism 2 function?	How did mechanism 2 generate process effects?	Agency comparisons
Externally designed interventions often feel imposed. SPIRIT aimed to enable agencies to identify local goals and tailor workshop content, but agencies were time pressured. Participants had extensive expertise in crafting policy, and many were experienced program designers and implementers using increasingly collaborative, bottom-up models. They viewed SPIRIT through this professional lens. The intervention's start date was randomised. CEOs decided if their agency would participate and nominated liaison people. Agencies encouraged participation but, as others note, participation in organisational interventions can feel 'expected'. ^b	Self-determination is the feeling of having control ^c . Interventions that foster self-determination share power and allow participants to pursue a variety of goals. Participants felt they had some control in relation to SPIRIT when there was: 1. <i>Flexibility</i> – scope to use the intervention to address their needs. 2. <i>Decision-making support</i> the audit feedback and deliberative processes helped leaders to make informed decisions, and those involved in tailoring received guidance about what could be achieved and how to do it. 3. <i>Locally shaped content</i> – managers and liaison people actively tailored goals and content, and ensured colleagues had a say in it (M9); interactive workshops enabled participants to drive content; and participants co-designed and co-presented workshops. 4. <i>Choice</i> about whether to take part in the intervention and data collection or be a liaison person, irrespective of managerial expectations. Self-determination had to be tempered with judicious decision-making. Some choices backfired such as when liaison people crammed content into workshops which overwhelmed participants.	Tailoring and interactivity were consistently viewed as critical for getting value out of participation. Self-determination encouraged leaders and liaison people to champion SPIRIT, and tailoring gave them a key selling point – “it's designed for us”. When self-determination was constrained (e.g. in didactic workshops) it tended to frustrate participants. Tailoring was time-consuming and sometimes required new decision-making pathways, so some agencies found it burdensome and did not make full use of the flexibility on offer (even though they were adamant it was necessary). Where liaison people rejected the premise of SPIRIT (M1) or believed that it was not flexible enough, they put less effort into tailoring and promotion. Leaders wanted choice about when SPIRIT started so it could be used more strategically to complement (or avoid) other activities.	Agencies 1,3, 5 and 6 felt they had enough scope to shape intervention content and did so moderately to extensively. Where there was scope, participants in all agencies contributed to workshop content, increasing its relevance and applicability. Greater involvement seemed to increase receptivity and investment in outcomes. Less effort was made to tailor the intervention in A2 and A4. This may have increased expectations of (and actual) incompatibility. Several liaison people were reluctant to take on their role, and many participants in A2 felt obliged to participate. Despite considerable encouragement, only A5 co-presented a workshop. Two agencies modified non-flexible aspects of SPIRIT: A3 insisted on different participant eligibility criteria to reduce the burden on staff, and A6 requested a hiatus in the intervention while they managed a restructure.

^b Aguinis H, Henle CA: *Ethics in research. In Handbook of research methods in industrial and organizational psychology. Edited by Rogelberg S. Oxford: Blackwell; 2004: 34-56*

^c Srivastava UR, Singh M: *Psychological empowerment at the work place. Global J Bus Man 2008, 2:53-73.*

Mechanism interactions and feedback

As others have noted, separating interactive processes into discrete mechanisms, while useful for theory building, fails to reflect their interdependence [61]. Many of the nine mechanisms include related concepts, which in some cases may be nested. For example, 'self-determination' (M2) is linked with 'respect' (M6) and may function as a mechanism within 'self-efficacy' (M5).

Figure 2 illustrates feedback within our model. This accords with the realist view that contexts, mechanisms and outcomes are not fixed entities but are contingent on the focus of the current evaluation, i.e. they function as a context, mechanism or outcome in a particular part

of the analysis. Thus, many of our process effects feed back into and overlap functionally with the identified mechanisms, and may well function as mechanisms when this data is combined with the study outcomes. This is especially pertinent in a process evaluation given that process effects are hypothesised to mediate the intervention outcomes. An example of feedback is our finding that 'persuasive leadership' is a mechanism, despite one of the process effects being 'Leaders support SPIRIT'. This is because we found 'persuasive leadership' to be crucial in activating other mechanisms (e.g. in asserting SPIRIT's value proposition) and thus in achieving many of the other process effects.

Table 4 Mechanism 3 – The value proposition

Proposition: Busy policymakers decide whether to participate based on the intervention's value proposition. This is determined by what is on offer, at what cost, and how it is communicated. Value can be identified in individual components even where aspect of the overall intervention or trial are rejected. Value is anticipated where the content promises to be useful, stimulating, aligned with local goals, and where there are clear answers about *what, why, who* and *when*. Agency-attuned communication is essential.

Context	How did mechanism 3 function?	How did mechanism 3 generate process effects?	Agency comparisons
<p>The ideal of research-informed policy was espoused in the wider environment, but each agency had a distinct organisational culture that interpreted this differently. Busy policy staff were juggling competing demands and needed a good reason to take part in non-essential activities. They calculated trade-offs: 'what can I afford to lose or postpone to make way for SPIRIT?' All suffered generalised information overload, but many complained about the lack of useful research in their area.</p> <p>The challenge of explaining SPIRIT was exacerbated by: 1. a complex and unfamiliar study design; 2. flexibility (it was being tailored and in flux); and 3. two levels of outcomes: those of the trial (fixed) and those identified by the agency (targeting local goals).</p>	<p>A "value proposition" (promised advantage) is a convincing argument about the worth of a strategy that is assessed by prospective users on the basis of perceived costs and benefits^d. Participants' view of SPIRIT's value proposition related to: 1. <i>Utility</i> - the content promised to be relevant and applicable, addressing current or future needs. Knowing SPIRIT was locally tailored increased expectations of utility. 2. <i>Stimulation</i> - content promised to be interesting. Presenters with "big names", expert roles, and very senior policy experience piqued interest. 3. <i>Persuasive marketing</i> – clear communication using agency-attuned language that emphasised the value of SPIRIT, framed it in relation to agency values and goals, and was disseminated through locally appropriate channels. 4. <i>Forecasting</i> – the perceived quality of each intervention activity was used as an indicator of the likely quality of further activities, but only where participants were aware that they were all part of SPIRIT. The value proposition differs from M1 in that it was assessed in relation to each activity - the premise of SPIRIT might be rejected but individual workshops could still promise value.</p>	<p>Where managers saw the <i>value proposition</i> they espoused SPIRIT and encouraged participation, appearing genuine in their efforts. Liaison people supported SPIRIT based on (a) the extent to which it seemed likely to benefit their agency, and (b) whether acting as the LP would benefit or disadvantage them personally. Where liaison people saw the value proposition they went the extra mile to ensure the agency benefitted. Where staff saw the value proposition, participation was moved higher in their list of priorities. The potential benefits of SPIRIT counteracted the burden of data collection ("<i>survey fatigue</i>") for agencies that had several data collection points prior to the intervention. Initially, SPIRIT marketing was suboptimal: dense, confusing, with poorly attuned "researchy" language (e.g. jargon and acronyms) – policymakers couldn't see the value proposition. Strategic advice and input from liaison people improved communications substantially.</p>	<p>Most interviewees in A1, A5 and A6 saw potential value in SPIRIT and so were receptive and inclined to participate. In A3, persuasive internal marketing increased the value proposition from a lower base. Some managers in all agencies encouraged their staff to participate based on potential value. A2 and A4 generally saw less potential value in SPIRIT, and their liaison people put less effort into shaping and promoting it. Some staff in all sites were confused about SPIRIT's purpose and form, did not know what was expected of them, and entangled the intervention with the trial. A particularly poorly received introductory session in A2 appeared to have lasting effects on perceptions of SPIRIT as a whole, despite some very well received workshops that followed.</p>

^d Barnes C, Blake H, Pinder D: *Creating and delivering your value proposition: managing customer experience for profit*. Kogan Page Publishers; 2009.

We also concluded that mechanisms functioned on a continuum that encompassed negative and positive expressions. Mechanisms were activated to different extents in each agency and, on occasion, were activated negatively. For example, several interviewees made it clear that mechanisms such as 'Self-determination', 'Getting good stuff' and 'Respect' were activated negatively when they were instructed by their manager to attend a 2-hour workshop that had no relevance to their work

Revised programme theory

These results enabled us to revise our programme theory to reflect contextual contingency, which also increases the

operational transferability to other interventions and settings (Table 11).

Discussion

From the participants' perspective, the most positive attributes of the intervention were useful (i.e. relevant and applicable) content, high profile experts who delivered pragmatic content and demonstrably "got it", active participation in intervention activities, and intervention flexibility supported by deliberative audit and feedback that informed goal-setting and customisation. Much of SPIRIT's implementation fidelity was sound – all the components of the intervention were

Table 5 Mechanism 4 – “Getting good stuff”

Proposition: Where policymakers want increased research knowledge, skills or resources, and an intervention offers useful, credible, tangible and stimulating content that facilitates participation, participants are likely to feel that they are *getting good stuff*. This encourages continued participation, internal promotion and identification of and intention to use ideas/resources. It also mitigates any data collection demands. Some participants can *get good stuff* irrespective of their view of the intervention/trial’s premise or broader value, but responses will be shaped by perceptions of current need.

Context	How did mechanism 4 function?	How did mechanism 4 generate process effects?	Agency comparisons
Policy work tackles complex problems in complex systems. Views about the value of research are contested. Some policy staff were not interested in new skills or ideas, but most wanted and actively sought them. Many were unaware of existing resources in their agencies (human and technical). Each agency had distinct learning norms. Time pressures and reactive practices limited opportunities for policymakers to engage with new ideas and reflect on their practice. SPIRIT demanded considerable amounts of data collection (mostly because of the trial).	Participants felt they got good stuff from SPIRIT in relation to: 1. <i>Utility</i> - content was relevant, applicable, pitched at the right level, partly because presenters had “ <i>done their homework</i> ”. 2. <i>Credibility</i> - presenters with content and practice expertise “ <i>got it</i> ”, i.e. they understood the constraints and opportunities of policymaking and the need for pragmatism. 3. <i>Tangibility</i> - targeted case examples and problem-solving activities made concepts concrete. 4. <i>Stimulation</i> - dynamic presenters captured interest and imagination e.g. via compelling behind-the-scenes anecdotes. 5. <i>Linkage</i> - interactive activities connected participants to external experts and existing internal resources, and forged intra-agency connections by alerting them to colleagues with expertise or shared work agendas. 6. <i>Learning congruence</i> - activities leveraged preferred learning styles. 7. <i>Reflective space</i> - workshops provided opportunities for critical thinking. 8. <i>Orientation</i> - new staff found that workshop discussions provided insights into how their colleagues view, access and use research.	The perceived ‘return on investment’ of participation encouraged (or discouraged) continued engagement in the intervention and data collection, positive word-of-mouth, and receptivity to SPIRIT’s ideas and resources. Feedback form data indicated very positive views, but several interviewees found some content irrelevant, inapplicable or boring and were less inclined to participate in other activities, including data collection. Some spoke negatively to colleagues about their experience, possibly influencing their decision to participate.	Very high numbers of participants across all agencies got good stuff from the majority of workshops. Feedback form results in all sites were extremely positive. Most interviewees engaged with some content, including those who did not anticipate value, and many identified “ <i>specific realisable benefits</i> ”. Some interviewees believed they would put ideas into practice, and a few had done so, especially newer staff for whom SPIRIT was “ <i>formative</i> ”. Interview data suggests fewer participants in A2, A3 and A4 got good stuff, but there are only minor differences in their feedback form data compared to other agencies (Additional file 2).

Mechanism 4 focuses on intervention workshops. Perceptions of other intervention components were harder to access. Many interviewees were vague about whether they had seen CEO emails espousing SPIRIT (M8), or if they were receiving weekly updates about resources in the online portal. Those who had accessed the portal said they found it helpful (albeit cumbersome to access due to password restrictions), but were not able to identify any specific use. We did not manage to interview those involved in trialling the commissioned research services in A3 and A4. In the other agencies, the response was mixed—only A5 and A6 were entirely happy about the final product and could identify ways that it would be used. Dissatisfaction with these services mainly appeared to be an artefact of the trial: agencies found it difficult to identify a useful service and topic when responding to an external timeframe. Several participating agencies that struggled to select and tailor their commissioned research service had a history of using the service previously with high levels of satisfaction, but those occasions had been agency-initiated and thus needs-driven.

delivered – but activities were not always as interactive or as participant-driven as intended. Authentic in-person leadership support and committed liaison people were vital mediators, while obstacles included confusion about the purpose of participation in SPIRIT, perceptions of poor alignment with agency practices or priorities, and feeling misunderstood or judged. Previous organisational change initiatives and archetypal views of researcher-policymaker relations sometimes appeared to underpin expectations and frame some of the concerns. The data collection demanded by the stepped wedge evaluation was onerous, and aspects of the trial were often entangled with participants’ perceptions of the

intervention. Like many others, we found that pre-existing positive relationships between the agency and those involved in designing and implementing the intervention had considerable facilitative effects [75–77]. In our case, they helped to activate mechanisms such as respect and confidence.

Implications for intervention improvement

Given their pivotal importance, greater upfront engagement with each agency’s leadership and the nominated liaison person would have been beneficial. Local tailoring and shared decision-making was essential, but challenging for both the agency and the intervention team. For example, it

Table 6 Mechanism 5 – Self-efficacy

Proposition: Where policymakers have some research use skills and are open to further development, self-efficacy can be bolstered through pragmatism, affirmation, modelling, strengths-based dialogue and participation. When this occurs, participants are likely to experience the intervention positively and feel confident and motivated about putting ideas into action.

Context	How did mechanism 5 function?	How did mechanism 5 generate process effects?	Agency comparisons
As analytic thinkers, many (but not all) participants were keen to critique their own, their programs' and their organisations' research engagement. Levels of confidence varied: audit survey data suggested the majority of staff lacked experience and confidence in accessing, appraising and using research, whereas 2/3 of interviewees said their skills were reasonable (many had research qualifications and/or experience) but, in most cases, could be improved. Some were apprehensive at the excessive standards of research use that SPIRIT might require. The quality of leadership feedback about staff and agency performance varied.	Self-efficacy refers to people's beliefs about their capability to perform tasks and achieve goals ^e . Those who feel they have the skills to put ideas and resources into action are more likely to adopt them ^f . Self-efficacy was activated by: 1. <i>Pragmatism</i> - presenters advocated realistic "good enough" goals which assured participants they could achieve acceptable practice standards in using research. 2. <i>Affirmation</i> - participants felt they were building on well-established capabilities partly due to strengths-based audit and mid-way feedback, plus sensitive facilitation in workshop activities, and leaders conveying confidence in their staff. 3. <i>Modelling</i> - high profile experts recounted 'war stories', successes and hard-won lessons that countered idealism and echoed messy local attempts to solve problems using research. 4. <i>Experiential learning</i> - trying out tools and systems increased understanding and confidence. 5. <i>Demonstrating expertise</i> - interactive activities enabled participants to contribute valuable local knowledge and skills to SPIRIT content.	This was one of the least tangible mechanisms but it appeared to have substantial effects on the extent to which targeted policymakers participated in, and were receptive to, the intervention. Many participants who felt they or their agency could use research better were encouraged (and sometimes pleasantly surprised) by SPIRIT content, "We can do this!". They tended to express enthusiasm for using the ideas and resources the encountered. Some participants who felt overwhelmed by technicalities or perceived high standards of using research indicated they would probably "leave it to someone else". Those with greater existing confidence appeared to contribute more actively to the content of SPIRIT activities, thus potentially reinforcing their 'capability status' in their own eyes and within the organisation.	More interviewees in A1, A5 and A6 than in the other agencies said they were encouraged by SPIRIT. In all agencies, the audit feedback largely supported managers' understanding of, and confidence in, their staff's capacity. A minority of participants with less confidence felt overwhelmed (e.g. a very highly rated evaluation workshop in A5 caused a few participants to see evaluation as outside their ability), whereas some others already had high self-efficacy (in all agencies, but possibly A3 in particular). Mid-intervention feedback that showed progress, and which may have supported organisational-level self-efficacy, was not always disseminated within the agency. A few interviewees in A1, A3 and A5 felt that agency-level participation in SPIRIT indicated that their CEO lacked faith in staff capabilities.

^e Bandura A: *Self-efficacy in changing societies*. New York Cambridge University Press; 1995.

^f Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O: *Diffusion of innovations in service organizations: systematic review and recommendations*. *Milbank Q* 2004, 82:581-629.

^g Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC: *Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science*. *Implement Sci* 2009, 4:1-15.

was often difficult for agencies to make strategic use of processes that they had not initiated such as trialling the services for commissioning research. Advice from agencies about how tailoring could be best supported in their context may have been beneficial, but the process of tailoring will always demand time and effort. This reflects the underpinning need for agency leaders to be committed to participation from the start.

Despite being selected for broad similarities, the six participating agencies had markedly different remits, practices and conceptualisations of evidence. SPIRIT's audit and feedback process was effective in developing a shared understanding of each agencies' current and desired

research use capabilities, but better understanding of their practice goals and values, and greater collaboration in designing the intervention and data collection instruments (which every agency desired) could have sharpened the meeting of minds about what was needed and how to address it. Understanding what participants think about intervention goals, and using their ideas about what should be done in order to achieve those goals, is usually critical for success [78].

As noted previously, the realist distinction between intervention activities and mechanisms is crucial for theory-driven evaluation, but it is equally crucial in the development of context-sensitive intervention design and

Table 7 Mechanism 6 – Respect

Proposition: Where participants feel that SPIRIT values them and their work, they are more likely to participate and to be receptive to intervention content. Inferred disrespect or criticism may cause resentment, disengagement and possible generation or reinforcement of negative views about researchers. If the intervention aims to (or is thought to) overhaul valued practices, participants will reject the intervention’s ideas. Closely connected with M1 and M2.

Context	How did mechanism 6 function?	How did mechanism 6 generate process effects?	Agency comparisons
<p>SPIRIT participants were experienced professionals with high level skills in information synthesis and analytical thinking. They were specialists in using knowledge strategically in policy processes. Each agency had some history of working with researchers, albeit in quite different ways. In many cases these relationships were regarded as productive, yet stereotypes of arrogant ‘ivory tower’ researchers were pervasive. Some staff had experience of researchers patronising them or treating policy work as unsophisticated. Some agencies had recent experiences of data collection with no feedback.</p>	<p>Targeted policymakers felt respected (or, at least, did not appear to feel disrespected) where they regarded SPIRIT as:</p> <ol style="list-style-type: none"> 1. <i>Contextually sensitive</i> - attuned to policymakers’ practices and environment (see M1). 2. <i>Strengths-orientated</i> – building on participants’ knowledge and skills. 3. <i>Supporting practice craft</i> - SPIRIT accommodated the local “<i>art of policymaking</i>” rather than attempting to impose work practices. 4. <i>Eliciting meaningful contributions</i> - participants’ expertise was valued and used in shaping content (M2). 5. <i>Responsive</i> - SPIRIT staff and presenters listened, adapted and responded constructively to participants’ queries, concerns and ideas. 6. <i>Providing feedback</i> - regular meaningful feedback was given. 7. <i>Respectful language</i> - communication and measures were succinct, non-patronising and strengths-orientated. 	<p><i>Respect</i> probably impacted all process effects, but especially no. 3: <i>Targeted policymakers participate in, and are receptive to, intervention activities</i>. Most participants appeared to feel understood and heard by the SPIRIT team, which supported engagement, willing participation and openness to what the intervention offered. The few interviewees who experienced aspects of SPIRIT as disrespectful were more critical of the intervention, and expressed more negative views about “<i>typical researchers</i>” and the challenges of researcher-policymaker relationships (countering a key intervention goal of enhancing these relationships). It seemed that these views were usually reinforced, rather than triggered, by their experience of SPIRIT.</p>	<p>The link between mutual positive regard and process effects was particularly noticeable in A1 where the liaison person responded to concerns on behalf of the SPIRIT team, and interviewees explained they felt more invested in SPIRIT as a consequence. Only a minority of participants indicated they felt disrespected, but we observed occurrences in all agencies except A5. It was due to: perceptions of a patronising or ill-informed premise, perceived assumptions in the measures or in how SPIRIT was introduced, lack of feedback and, on a few occasions, because of a presenter’s comments or delivery style. Belief that the SPIRIT team had made insufficient efforts to understand their needs, practices or context seemed more pronounced in policymakers who did not know the SPIRIT team.</p>

implementation planning. An intervention cannot simply ‘do’ respect, or ‘deliver’ self-efficacy, it cannot control the perceived attractiveness of its premise, or make internal facilitators act strategically. Activating these mechanisms is an evolving work-in-progress shaped by personalities, relationships and complex shifting environmental opportunities and constraints. Greater understanding of the mechanisms that generate desired (and undesired) process effects provides helpful guidance, but putting this learning into practice takes creativity, humility and reflexivity.

Our contribution

These findings add to the existing knowledge by surfacing evidence about how policymakers perceived and engaged with different aspects of an intervention trial designed to increase the extent to which they use research in their work. Our realist process evaluation approach goes beyond questions of implementation fidelity and ‘what works?’ to provide a more nuanced and theoretically informed account of how the intervention

produced process effects, and why there was such variation across the six policy agencies.

As per Fig. 2, we anticipate that the intervention’s process effects, and the mechanisms that underpin them, mediate the study outcomes, but we caution against assumptions that this is a linear predictive relationship. As realist evaluation adherents indicate, there are usually multiple causal pathways in real world interventions, and the best we can do is identify common pathways for particular groups of individuals in particular circumstances; therefore, we concur with McMullen et al. that, “*there is not, nor can there ever be, a universal implementation model for complex interventions. Site-specific characteristics and realities need to be considered*” [79]. However, this consideration need not start from scratch with each new intervention – we can develop an increasingly sophisticated understanding of the conditions that make these outcomes more likely in a given setting. As Pawson argues, “*evaluation science assumes that there will be some pattern to success and failure across*

Table 8 Mechanism 7 – Confidence

Proposition: Depending on their concerns—which may be affected by data fatigue, previous initiatives, professional interest in the intervention design, and scepticism about researchers—policymakers want to know that a trial in which they participate is scientifically and ethically trustworthy, and that participation is meaningful and poses no threat. Lack of confidence can lead to poor internal facilitation and damage relationships with researchers. Trial activities will affect perceptions of the intervention.

Context	How did mechanism 7 function?	How did mechanism 7 generate process effects?	Agency comparisons
<p>The trial required a lot of data collection: six measurement points using three measures, and a process evaluation. Local goal-setting was informed by audit findings from two of the measures. All agencies had previously endured disruptive change initiatives which, in some cases, had little perceived benefit. Many participants had considerable expertise in intervention research and evaluation. Some felt researchers were naïve about policymaking and there was a degree of scepticism about what a researcher-initiated intervention could offer.</p>	<p>Participants had confidence in SPIRIT when they regarded the intervention, and the trial, as: 1. <i>Valid</i> - (a) the audit and feedback data seemed robust thus (b) the goal-setting was well founded and (c) subsequent data collection promised to track meaningful change and provide useful findings. 2. <i>Trustworthy</i> - the SPIRIT team were seen to act transparently and non-judgementally, in good faith. Where there were positive pre-existing relationships between agency and SPIRIT staff it facilitated trust in the initial stages. 3. <i>Transparent</i> - it was clear what demands would be made on participants and how data would be used, including when and how outcomes would be communicated. 4. <i>Safe</i> - it was clear that reporting the trial would not compromise individuals or agencies. 5. <i>Effective</i> – the intervention strategies could generate meaningful change in their setting.</p>	<p><i>Confidence</i> affected the extent to which targeted policymakers wanted to take part in intervention activities and data collection. Liaison people reported that the audit feedback increased leaders' confidence in, and enthusiasm for, SPIRIT which they communicated to their staff. Participants who lacked confidence in the measures or intervention design, and the small minority who questioned the study's integrity or safety, expressed discomfort about SPIRIT. Some were avoidant and spoke poorly of researchers. Negative views of the trial were often entangled with the intervention. It seemed that scepticism about researcher-developed content seldom prevented participation, and follow up interview data suggested that attendance at a workshop tended to increase confidence.</p>	<p>45 leaders took part in audit feedback, and 38 completed evaluation forms. 37 of these (the exception was in A3) answered 'Yes' to the following statements: 1. <i>The forum provided clear and accessible information.</i> 2. <i>It provided useful feedback on how we currently use research.</i> 3. <i>The presenter had appropriate knowledge and skills.</i> 4. <i>It gave me confidence that SPIRIT will be tailored to suit this agency.</i> 5. <i>I will encourage my staff to participate in SPIRIT.</i> Despite this high level of confidence, information about the audit was often not disseminated effectively throughout agencies, thus many interviewees did not understand how it had been interpreted or how subsequent data would be used. In A1, a small minority of interviewees initially felt SPIRIT might threaten them or their agency with exposure, but some in all agencies expressed discomfort in "being researched". Some in A2 questioned the integrity of the study, suggesting it was a business endeavour disguised as research. Liaison people and leaders were instrumental in addressing concerns and increasing confidence, but in A4 and A6 the liaison people themselves questioned the sensitivity of the measurement instruments.</p>

interventions, and that we can build a model to explain it" [1]. We hope to have made a start in identifying these patterns in a form that will enable others to extrapolate and apply lessons to other interventions and contexts [1].

Strengths and limitations of this process evaluation

Using a realist approach enabled us to identify and test hypothesised causal mechanisms, evaluate the extent to which SPIRIT activated them, use this analysis to refine the programme theory, and identify areas of strength and potential improvement in the intervention and trial design. The identification of underlying causal mechanisms and the development of propositions enhances the utility and transferability of the findings [3, 80] and strengthens the general knowledge base by building on existing theories. The thematic overview of the process evaluation data in Additional file 1, and the inclusion of informing theory in

Additional file 2, provide 'analytical trails' that support the findings.

Triangulating different types of data obliged us to consider diverse points of view and increased the trustworthiness of our findings. As Wells et al. [9] note, "... evaluations need to incorporate multiple methods, multiple sources and multiple perspectives if they are to reflect the context of practice adequately". We achieved this thanks to (1) the unusually generous appointment of a dedicated process evaluation researcher throughout the study, and (2) the length of the intervention (12 months) and its staggered delivery, which gave us considerable time in each agency to test hypotheses at different points in the intervention across six sites. However, we acknowledge this was an exploratory first step and the ideas are yet to be tested by others and in different settings; therefore, at this stage, our findings are only a rough indication of

Table 9 Mechanism 8 – Persuasive leadership

Proposition: Leaders persuade policymakers they believe in SPIRIT where the leaders: have credibility as research advocates; support intervention goals visibly, consistently and authentically; articulate the intervention's value; and model engaged participation. Messages about the need for change must be balanced with assurance of existing capabilities. Expert presenters can be persuasive and inspiring leaders who model values in the wider system.

Context	How did mechanism 8 function?	How did mechanism 8 generate process effects?	Agency comparisons
Participating agencies were bureaucracies with strict hierarchies, but very different infrastructures and numbers of staff. Policymakers in the six agencies had varying levels of respect for their leaders, including different views about the extent to which their leaders valued using research. The professional behaviours of agency leaders in relation to SPIRIT were interpreted locally.	Leaders (managers, opinion leaders and liaison people) were persuasive in support of SPIRIT when: 1. they engaged in <i>value messaging</i> , i.e. they articulated the benefits of SPIRIT including their agency's <i>scope for and need for</i> increased capacity in using research (M1) in a manner that did not demean current capacity (M5). 2. Respected managers <i>modelled engagement</i> with the interventions' ideas and resources, thereby connecting SPIRIT to organisational values and priorities. When impressive expert presenters modelled their commitment to research-informed policy and provided examples of its benefits they positioned SPIRIT's values in the wider policy environment. 3. <i>Agency leadership support was: visible; credible</i> (these managers were known to support and engage in research-informed policymaking themselves); <i>consistent</i> (leaders across the agency conveyed support); and <i>authentic</i> (support seemed genuine). In-person advocacy and participation in workshops was experienced as more authentic than email espousal - actions speak louder than words.	<i>Persuasive leadership</i> probably impacted all the other mechanisms and process effects, and was a process effect in its own right (see section on mechanism interactions and feedback below). For example, it connected SPIRIT to organisational priorities (M1), increased perceptions of potential value (M3) and trustworthiness (M7), gave staff a positive message about the relationship between SPIRIT and current capabilities (M5), and provided liaison people with a mandate for action (M9). All these impacts, including expert presenters' advocacy for pragmatic research-informed policy, increased receptivity to the intervention.	Leadership support was strongest in A1 and A5. In other agencies managers either lacked credibility as research advocates or their espousal appeared to be perfunctory or inconsistent (e.g. the CEO supported SPIRIT but some managers voiced dissent). Some presenters had insufficient policy-savvy credibility to function as leaders – in all agencies to a small extent, but it was 'felt' more in A2 and A4 where resistance was higher. In both A3 and A5 managers' attendance in workshops bolstered perceptions of SPIRIT's importance but also inhibited frank discussion by more junior staff. This may have occurred to some extent in other agencies. Forums targeting agency leaders were especially well attended when scheduled within or in lieu of formal management meetings.

major causal patterns within SPIRIT's engagement and participation. Further testing and refinement are required.

A limitation was our inability to determine the full range of views and experiences of targeted staff in each agency. Interviewees were sampled purposively for maximum variation of relevant views and experiences, but many declined interviews and it was not always possible to identify substitutes. Others have found similar problems [52]. Consequently, we reached a smaller range of participants than envisaged and so may have missed important views. For example, all the process evaluation interviewees in A4 (11 people with a total of 15 interviews over the duration of the intervention) were either lukewarm or dismissive of SPIRIT, but during outcome measures interviews some A4 participants stated that they welcomed the intervention, and following the trial their CEO said SPIRIT had impacted his agency positively. In all agencies, we saw some non-agreement between the highly positive feedback form data and the more critical responses in the interview data. This may be the result

of different foci – interviews ranged across the whole of SPIRIT (including its premise, communication and data collection), while feedback forms were workshop-specific – but other factors could be skewed sampling, leading interview questions or the bluntness of the feedback form. The response rate for feedback forms was good, with 74% of attendees completing them, but it is unclear whether those who did not complete forms differed from those who did, and thus what views we might have missed. The direction of this quantitative data was consistent with patterns in the qualitative data regarding a more positive response from agencies 1, 5 and 6, but feedback form responses across agencies and items were so similar that it is likely that the tool discriminated poorly. We used Yes/No statements to maximise response rates from participants who might be rushing to leave, but this was probably too limiting. Certainly, there were many occasions where the free text fields conveyed ambivalence or, at least, scope for improvement, when the scored statements suggested 100% satisfaction. We would use a more sensitive instrument in the future.

Table 10 Mechanism 9 – Strategic insider facilitation

Proposition: Liaison people are internal staff appointed to coordinate SPIRIT. Where they actively use strategic insider knowledge to support access to and uptake of the intervention via translation, mediation and locally appropriate promotional strategies, they hugely increase awareness and understanding of the intervention, encouraging greater receptivity and participation. Managerial support is required and, even more than other staff, liaison people must anticipate potential value from SPIRIT.

Context	How did mechanism 9 function?	How did mechanism 9 generate process effects?	Agency comparisons
<p>SPIRIT required that each CEO nominate a local 'liaison person' to coordinate SPIRIT in their agency. Liaison people and managers had their own views about using research in policy work, and about the value of SPIRIT's goals and strategies. In most cases they did not have a say in the agency-level decision to participate in SPIRIT. Beliefs about divisions between the 'worlds' of research and policy had currency.</p>	<p>In order to be strategic and facilitative, the internal coordination of SPIRIT depended on: 1. <i>Translation</i> - liaison people and managers used their 'insider' expertise to explain SPIRIT in terms that made most sense to targeted policymakers, illustrating how the intervention intersected with and complemented organisational goals and activities. 2. <i>Mediation</i> - liaison people actively identified concerns and worked with the SPIRIT team to resolve them. 3. <i>Persuasive marketing</i> - liaison people used local communication channels and creative strategies to "sell" SPIRIT (see M3). 4. <i>Negotiation</i> - liaison people and, in one agency, managers, advocated forcefully for adaptations to the intervention that would better suit their needs 5. <i>Support</i> - liaison people were supported internally by managers and externally by a responsive SPIRIT team who provided materials, information and feedback (M6).</p>	<p>Like M8 above, strategic insider facilitation strongly affected other mechanisms as well as SPIRIT's process effects. High quality facilitation fostered consultative tailoring, continuous engagement and informed participation. Information reached targeted staff in a form that was persuasive and accessible. Problems were identified and resolved so there was greater confidence and receptivity. Poor facilitation led to suboptimal tailoring, confusion about the intervention's purpose and form, and a poorer local value proposition (M3), resulting in lower levels of participation or unwilling attendance.</p>	<p>Liaison people in A1, A3, A5 and A6 used creative strategies to tailor and champion SPIRIT, disseminating information through formal and informal channels, e.g. by nominating colleagues to give updates at their team meetings and promoting forthcoming events over coffee. There were transitional lags during changes of liaison person in A3, A4 and A6. The appointment of an enthusiastic and well-supported liaison person in A3 increased survey response rates. Most liaison people spoke highly of the SPIRIT team's support, but A6 experienced too many points of contact and A4 said there was insufficient guidance. As noted, the liaison people in A2 and A4 did not facilitate SPIRIT to the best of their abilities because they did not believe it was worthwhile (M1, M3, M7). This probably contributed to a lower value proposition (M3) for targeted participants in those agencies, and less confidence (M7). The perceptions, behaviour and impact of liaison people is covered in more detail elsewhere ^h.</p>

^h Haynes A, Butow P, Brennan S, Williamson A, Redman S, Carter S, Gallego G, Rudge S: *The pivotal position of 'liaison people': facilitating a research utilisation intervention in policy agencies. Evid Policy 2016 (online 14 Dec).*

Table 11 Initial and revised programme theory

Initial programme theory (a-contextual)	Revised programme theory (contextually contingent)
<p>SPIRIT will engage and motivate agency leaders to 'own' the intervention using audit feedback, deliberative goal-setting and programme tailoring –this agency-driven approach will generate a priority-focused programme that offers locally relevant practice support and accommodates differences in agencies' values, goals, resources and remits. The programme will comprise a suite of andragogical activities, tools and connections across the research-policy divide that provide resources and build knowledge, skills and relationships, and will be supported via modelling and opinion leadership by agency leaders and dynamic external experts. CEOs will promote SPIRIT in their agencies and liaison people will facilitate the tailoring and implementation – these strategies will act synergistically to stimulate and resource participants at different organisational levels, leading to changes in values, practice behaviours and agency processes. This will facilitate increased use of research in policy processes</p>	<p>Where agencies have an existing orientation to use academic research and are on a trajectory of improved use with perceived room for improvement, SPIRIT will be used to complement or trigger organisational initiatives. Where liaison people and agency leaders believe in the value of the intervention and have confidence in the measures, they will play a pivotal role in tailoring the intervention and championing its goals. Leaders will be motivated by deliberative audit feedback and goal-setting. In all sites, ownership will be increased by greater consultation, collaboration and choice. Agency-attuned communications will be vital in explaining goals, conveying value and addressing concerns. Andragogical activities, tools and connection across the research-policy divide will be valued in all agencies where they leverage existing strengths and address local concerns pragmatically. Staff will make use of these opportunities where they see concrete benefits, and newer staff may benefit most</p>

Reflections on conducting a realist process evaluation

Conducting a realist process evaluation was immensely valuable, but time consuming and challenging. Like others (e.g. [49, 81]), we struggled to disentangle aspects of the causal pathways; specifically, to delineate mechanisms from intervention strategies, contexts and outcomes. Realist analysis does not have a step-by-step guide, and it presents a unique tension between ontology and epistemology, so we sometimes struggled to reconcile our search for factual existing mechanisms with the need to take an “*imaginative leap*” and postulate those mechanisms [82]. Three strategies helped: first, scanning appropriate literature and drawing on established theories, for example, the concept of relative advantage [6, 58, 83] was critical for understanding variation in perceptions of SPIRIT and how this linked to the communication strategy. Second, the realist emphasis on counterfactual thinking [54] was very helpful in weighing up the plausibility of different theories. Third, reminding ourselves that causality does not function as discrete components or configurations and that our analysis was intentionally abstracting for the purposes of theory building rather than attempting to depict reality in all its messy, interdependent glory (see also [61]).

Conclusion

This realist process evaluation describes how participants experienced different aspects of a multi-component research utilisation intervention in policy organisations, and why there was such variation across the six implementation sites. We identify nine mechanisms that appeared to facilitate engagement with and participation in the intervention in these settings: (1) Accepting the premise (agreeing with the study’s assumptions), (2) Self-determination (participative choice), (3) The value proposition (seeing potential gain), (4) ‘Getting good stuff’ (identifying useful ideas, resources or connections), (5) Self-efficacy (believing ‘we can do this!’), (6) Respect (feeling that SPIRIT understands and values one’s work), (7) Confidence (believing in the study’s integrity and validity), (8) Persuasive leadership (authentic and compelling managerial advocacy) and (9) Strategic insider facilitation (local translation and mediation). This analysis was used to develop tentative propositions and to revise the overarching programme theory. Although our findings are nascent and require further testing and refinement, they indicate areas of strength and weaknesses that can guide the development and implementation of similar studies in other settings, increasing their sensitivity to the range of issues that affect the value and compatibility of interventions in policy agencies.

Additional files

Additional file 1: Descriptive overview of results. (PDF 403 kb)

Additional file 2: Supporting theory. (PDF 335 kb)

Additional file 3: Summary of SPIRIT intervention implementation fidelity. (PDF 1460 kb)

Abbreviations

CEO: Chief executive officer; SPIRIT: Supporting Policy In health with Research: an Intervention Trial

Acknowledgements

Many thanks to the agencies and individuals who participated in SPIRIT. We are indebted to the policymakers who took time out of busy schedules to contribute to our process evaluation. Thanks also to Emma Darsana for help with data management and coding; to Huw Davies, Stacy Carter and Joanne McKenzie, who offered advice on previous versions of this paper; and to the reviewers who gave us constructive feedback.

Funding

SPIRIT was funded as part of the Centre for Informing Policy in Health with Evidence from Research CIPHER), an Australian National Health and Medical Research Council (NHMRC) Centre for Research Excellence (#1001436), administered by the Sax Institute. CIPHER is a joint project of the Sax Institute; Australasian Cochrane Centre, Monash University; University of Newcastle; University of New South Wales; Research Unit for Research Utilisation, University of St Andrews and University of Edinburgh; Australian National University; and University of South Australia. The Sax Institute receives a grant from the NSW Ministry of Health. The Australasian Cochrane Centre is funded by the Australian Government through the NHMRC. AH is supported by an NHMRC Public Health and Health Services Postgraduate Research Scholarship (#1093096).

Authors’ contributions

AH led the process evaluation design, data collection and analysis, and drafted the manuscript. SB contributed to the design, independently conducted some analyses and, together with AH, workshopped initial findings to refine them. SM led the statistical analyses. All authors made substantial contributions to the analysis and interpretation of data, and were involved in critically revising the manuscript for important intellectual content. All authors read and approved the final manuscript.

Authors information

Not applicable.

Ethics approval and consent to participate

All participants provided informed consent. Both the agencies and individual staff members were advised they were free to decline to participate in any or all aspects of SPIRIT at any time without explanation. Ethical approval was granted by the University of Western Sydney Human Research Ethics Committee, approval numbers H8855 and H9870.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Sax Institute, 235 Jones Street, Ultimo, NSW 2007, Australia. ²Sydney School of Public Health, Edward Ford Building (A27), University of Sydney, Camperdown, NSW 2006, Australia. ³Australasian Cochrane Centre, School of Public Health and Preventive Medicine, Monash University, 553 St Kilda Road, Melbourne, Victoria 3004, Australia. ⁴School of Medicine, University of Notre Dame, 160 Oxford St, Darlinghurst, NSW 2010, Australia. ⁵Centre for Medical

Psychology & Evidence-based Decision-making, University of Sydney, The Lifehouse, 119-143 Missenden Rd, Camperdown, NSW 2006, Australia.

Received: 19 April 2017 Accepted: 18 July 2017

Published online: 23 November 2017

References

- Pawson R. *The Science of Evaluation: A Realist Manifesto*. London: Sage; 2013.
- Greenhalgh T, Wong G, Jagosh J, Greenhalgh J, Manzano A, Westhorp G, Pawson R. Protocol—the RAMESES II study: developing guidance and reporting standards for realist evaluation. *BMJ Open*. 2015;5(8):e008567.
- Pawson R, Tilley N. *Realist Evaluation*. South Australia: Community Matters; 2004.
- Astbury B. Some reflections on Pawson's Science of evaluation: a realist manifesto. *Evaluation*. 2013;19:383–401.
- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4:50.
- Funnell SC, Rogers PJ. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*. San Francisco: Jossey-Bass; 2011.
- Ritchie J, Lewis J. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE; 2003.
- Rycroft-Malone J, Fontenla M, Bick D, Seers K. A realistic evaluation: the case of protocol-based care. *Implement Sci*. 2010;5:38.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, Moore L, O'Cathain A, Tinati T, Wight D, Baird J. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.
- Astbury B, Leeuw FL. Unpacking black boxes: mechanisms and theory building in evaluation. *Am J Eval*. 2010;31:363–81.
- Nielsen K. How can we make organizational interventions work? Employees and line managers as actively crafting interventions. *Hum Relat*. 2013;66:1029–50.
- Nielsen K, Randall R. Opening the black box: Presenting a model for evaluating organizational-level interventions. *Eur J Work Organ Psy*. 2013;22:601–17.
- Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q*. 2004;82:581–629.
- McLean S, Moss G. They're happy, but did they make a difference? Applying Kirkpatrick's framework to the evaluation of a national leadership program. *Can J Prog Eval*. 2003;18:1–23.
- Wells M, Williams B, Treweek S, Coyle J, Taylor J. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials*. 2012;13:95–111.
- Datta A, Shaxson L, Pellini A. *Capacity, Complexity and Consulting: Lessons from Managing Capacity Development Projects*. ODI Working Paper 344. London: Overseas Development Institute; 2012.
- Moore GF. *Developing a Mixed Methods Framework for Process Evaluations of Complex Interventions: The Case of the National Exercise Referral Scheme Policy Trial in Wales*. Cardiff: Cardiff University; 2010.
- Schein EH. *Organizational Culture and Leadership*. London: John Wiley & Sons; 2010.
- Devos G, Buelens M, Bouckennooghe D. Contribution of content, context, and process to understanding openness to organizational change: two experimental stimulation studies. *J Soc Psychol*. 2007;147:607–30.
- Dixon-Woods M, Leslie M, Tarrant C, Bion J. Explaining Matching Michigan: an ethnographic study of a patient safety program. *Implement Sci*. 2013;8:70.
- Andersen H, Røvik KA. Lost in translation: a case-study of the travel of lean thinking in a hospital. *BMC Health Serv Res*. 2015;15:401.
- Fixsen DL, Naoom SF, Blase KA, Friedman RM, Wallace F. *Implementation Research: A Synthesis of the Literature*. Florida Mental Health Institute Publication 231. Tampa: National Implementation Research Network, University of South Florida; 2005.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008;337:a1655.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, Moore L, O'Cathain A, Tinati T, Wight D, Baird J. *Process Evaluation of Complex Interventions: Medical Research Council Guidance. A Report Prepared on Behalf of the MRC Population Health Science Research Network*. London: Institute of Education; 2015.
- Hodinnott P, Britten J, Pill R. Why do interventions work in some places and not others: A breastfeeding support group trial. *Soc Sci Med*. 2010;70:769–78.
- Ferlie E, Dopson S, Fitzgerald L, Locoock L. Renewing policy to support evidence-based health care. *Public Adm*. 2009;87:837–52.
- Oliver K, Lorenc T, Innvæ S. New directions in evidence-based policy research: a critical analysis of the literature. *Health Res Policy Sys*. 2014;12:34.
- Haynes A, Gillespie JA, Derrick GE, Hall WD, Redman S, Chapman S, Sturk H. Galvanizers, guides, champions, and shields: the many ways that policymakers use public health researchers. *Milbank Q*. 2011;89:564–98.
- Liverani M, Hawkins B, Parkhurst JO. Political and institutional influences on the use of evidence in public health policy. A systematic review. *PLoS ONE*. 2013;8:e77404.
- Russell J, Greenhalgh T, Byrne E, McDonnell J. Recognizing rhetoric in health care policy analysis. *J Health Serv Res Policy*. 2008;13:40–6.
- Majone G. *Evidence, argument, and persuasion in the policy process*. New Haven, CT: Yale University Press; 1989.
- Rycroft-Malone J. The PARIHS framework—a framework for guiding the implementation of evidence-based practice. *J Nurs Care Qual*. 2004;19:297–304.
- Crilly T, Jashapara A, Ferlie E. *Research Utilisation & Knowledge Mobilisation: A Scoping Review of the Literature*. Report for the National Institute for Health Research Service Delivery and Organization. London: Queen's Printer and Controller of HMSO; 2010.
- Tomoaia-Cotisel A, Scammon DL, Waitzman NJ, Cronholm PF, Halladay JR, Driscoll DL, Solberg LI, Hsu C, Tai-Seale M, Hiratsuka V, et al. Context matters: the experience of 14 research teams in systematically reporting contextual factors important for practice change. *Ann Fam Med*. 2013;11:S115–23.
- Clark AM, MacIntyre PD, Cruickshank J. A critical realist approach to understanding and evaluating heart health programmes. *Health*. 2007;11:513–39.
- The CIPHER Investigators. Supporting Policy In health with Research: an Intervention Trial (SPIRIT)—protocol for a stepped wedge trial. *BMJ Open*. 2014;4(7):e005293.
- Makkar SR, Turner T, Williamson A, Louviere J, Redman S, Haynes A, Green S, Brennan S. The development of ORACLE: a measure of an organisation's capacity to engage in evidence-informed health policy. *Health Res Policy Syst*. 2016;14:4.
- Makkar SR, Brennan S, Turner T, Williamson A, Redman S, Green S. The development of SAGE: a tool to evaluate how policymakers' engage with and use research in health policymaking. *Res Evaluat*. 2016;25:315–28.
- Makkar SR, Williamson A, Turner T, Redman S, Louviere J. Using conjoint analysis to develop a system of scoring policymakers' use of research in policy and program development. *Health Res Policy Syst*. 2015;13:35.
- Makkar SR, Williamson A, Turner T, Redman S, Louviere J. Using conjoint analysis to develop a system to score research engagement actions by health decision makers. *Health Res Policy Syst*. 2015;13:22.
- Brennan SE, McKenzie JE, Turner T, Redman S, Makkar S, Williamson A, Haynes A, Green SE. Development and validation of SEER (Seeking, Engaging with and Evaluating Research): a measure of policymakers' capacity to engage with and use research. *Health Res Policy Syst*. 2017;15:1.
- Redman S, Turner T, Davies H, Williamson A, Haynes A, Brennan S, Milat A, O'Connor D, Blyth F, Jorm L, Green S. The SPIRIT Action Framework: A structured approach to selecting and testing strategies to increase the use of research in policy. *Soc Sci Med*. 2015;136–137:147–55.
- Fridrich A, Jenny GJ, Bauer GF. The context, process, and outcome evaluation model for organisational health interventions. *Biomed Res Int*. 2015;2015:Article ID 414832.
- Westhorp G, Prins E, Kusters C, Hultink M, Guijt I, Brouwers J. *Realist Evaluation: An Overview*. Report from an Expert Seminar with Dr Gill Westhorp. Wageningen: Centre for Development Innovation, Wageningen University; 2011.
- Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. *BMC Med*. 2016;14:96.
- Haynes A, Brennan S, Carter S, O'Connor D, Huckel Schneider C, Turner T, Gallego G. Protocol for the process evaluation of a complex intervention designed to increase the use of research in health policy and program organisations (the SPIRIT study). *Implement Sci*. 2014;9:113.
- Martin GP, Ward V, Hendy J, Rowley E, Nancarrow S, Heaton J, Britten N, Fielden S, Ariss S. The challenges of evaluating large-scale, multi-partner programmes: the case of NIHR CLAHRCs. *Evid Policy*. 2011;7:489–509.
- Marchal B, Dedzo M, Kegels G. A realist evaluation of the management of a well-performing regional hospital in Ghana. *BMC Health Serv Res*. 2010;10:24.
- Salter KL, Kothari A. Using realist evaluation to open the black box of knowledge translation: a state-of-the-art review. *Implement Sci*. 2014;9:115.

50. Pawson R. Middle range theory and program theory evaluation: from provenance to practice. In: Leeuw F, Vaessen J, editors. *Mind the Gap: Perspectives on Policy Evaluation and the Social Sciences*, vol. 16. Piscataway, NJ: Transaction Press; 2009. p. 171–203.
51. McEvoy P, Richards D. A critical realist rationale for using a combination of quantitative and qualitative methods. *J Res Nurs*. 2006;11:66–78.
52. Dalkin SM, Greenhalgh J, Jones D, Cunningham B, Lhussier M. What's in a mechanism? Development of a key concept in realist evaluation. *Implement Sci*. 2015;10:49.
53. Cartwright N. Knowing what we are talking about: why evidence doesn't always travel. *Evid Policy*. 2013;9:97–112.
54. Merton RK. On sociological theories of the middle range. In: *Social Theory and Social Structure*. New York: Simon & Schuster, The Free Press; 1949. pp. 39–53.
55. Puntun M, Vogel I, Lloyd R. Reflections from a realist evaluation in progress: scaling ladders and stitching theory. In: *CDI Practice Papers*. Brighton: Institute of Development Studies; 2016.
56. Abimbola S, Molemodile SK, Okonkwo OA, Negin J, Jan S, Martiniuk AL. The government cannot do it all alone: realist analysis of the minutes of community health committee meetings in Nigeria. *Health Policy Plan*. 2016;31(3):332–45.
57. Adams A, Sedalia S, McNab S, Sarker M. Lessons learned in using realist evaluation to assess maternal and newborn health programming in rural Bangladesh. *Health Policy Plan*. 2016;31:267–75.
58. Rushmer R, Hunter D, Steven A. Using interactive workshops to prompt knowledge exchange: a realist evaluation of a knowledge to action initiative. *Public Health*. 2014;128(6):552–60.
59. Haynes A, Brennan S, Redman S, Williamson A, Gallego G, Butow P. Figuring out fidelity: a worked example of the methods used to identify, critique and revise the essential elements of a contextualised intervention in health policy agencies. *Implement Sci*. 2016;11:23.
60. Danermark B, Ekstrom M, Jakobsen L, Karlsson J. *Explaining Society: An Introduction to Critical Realism in the Social Sciences*. London: Routledge; 2002.
61. Chen H-T. *Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness*. Thousand Oaks: Sage; 2005.
62. Kitson AL, Rycroft-Malone J, Harvey G, McCormack B, Seers K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARIHS framework: theoretical and practical challenges. *Implement Sci*. 2008;3:1.
63. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol*. 2009, 43:267–76.
64. Damschroder LJ, Lowery JC. Evaluation of a large-scale weight management program using the consolidated framework for implementation research (CFIR). *Implement Sci*. 2013;8:51.
65. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology*. 2009;43:535–54.
66. Normalization Process Theory Online Users' Manual, Toolkit and NoMAD Instrument. <http://www.normalizationprocess.org>. Accessed 3 May 2017.
67. Greenhalgh T, Humphrey C, Hughes J, Macfarlane F, Butler C, Pawson R. How do you modernize a health service? A realist evaluation of whole-scale transformation in London. *Millbank Q*. 2009;87:391–416.
68. Sayer A. *Realism and Social Science*. 2nd ed. London: Sage; 2000.
69. Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In: Ritchie J, Lewis J, editors. *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. London: SAGE; 2003. p. 219–62.
70. International Pty Ltd QSR. NVivo Qualitative Data Analysis Software: Version 10. 2012.
71. Charlson FJ, Moran AE, Freedman G, Norman RE, Stapelberg NJ, Baxter AJ, Vos T, Whiteford HA. The contribution of major depression to the global burden of ischemic heart disease: a comparative risk assessment. *BMC Med*. 2013;11:250.
72. Thorne S. *Interpretive Description*. Walnut Creek: Left Coast Press; 2008.
73. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. *Am J Eval*. 2006;27:237–46.
74. Silverman D. *Doing Qualitative Research: A Practical Handbook*. 4th ed. London: SAGE Publications; 2013.
75. Sayer A. *Method in Social Science: A Realist Approach*. 2nd ed. Abingdon: Routledge; 2010.
76. Sandelowski M. Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Res Nurs Health*. 2000;23:246–55.
77. Caracelli VJ, Greene JC. Data analysis strategies for mixed-method evaluation designs. *Educ Eval Policy An*. 1993;15:195–207.
78. Tremblay D, Touati N, Roberge D, Denis J-L, Turcotte A, Samson B. Conditions for production of interdisciplinary teamwork outcomes in oncology teams: protocol for a realist evaluation. *Implement Sci*. 2014;9:76.
79. Reed M. Reflections on the 'realist turn' in organization and management studies. *J Manage Stud*. 2005;42:1621–44.
80. Jagosh J, Pluye P, Wong G, Cargo M, Salsberg J, Bush PL, Herbert CP, Green LW, Greenhalgh T, Macaulay AC. Critical reflections on realist review: insights from customizing the methodology to the needs of participatory research assessment. *Res Synth Methods*. 2014;5:131–41.
81. Materially Social. Causality, Method and Imagination. 2016. <http://materiallysocial.blogspot.com.au/2016/05/causality-method-and-imagination.html>. Accessed 3 May 2017.
82. Dearing JW. Applying diffusion of innovation theory to intervention development. *Res Social Work Prac*. 2009;19:503–18.
83. Edwards M, Evans M. Getting evidence into policy-making: parliamentary triangle seminar report. In: *ANZSIG Insights*. Canberra: ANZOG Institute for Governance, University of Canberra; 2011.
84. Aguinis H, Henle CA. Ethics in research. In: Rogelberg S, editor. *Handbook of Research Methods in Industrial and Organizational Psychology*. Oxford: Blackwell; 2004. p. 34–56.
85. Srivastava UR, Singh M. Psychological empowerment at the work place. *Global J Bus Man*. 2008;2:53–73.
86. Barnes C, Blake H, Pinder D. *Creating and Delivering your Value Proposition: Managing Customer Experience for Profit*. London: Kogan Page Publishers; 2009.
87. Bandura A. *Self-efficacy in Changing Societies*. New York: Cambridge University Press; 1995.
88. Haynes A, Butow P, Brennan S, Williamson A, Redman S, Carter S, Gallego G, Rudge S. The pivotal position of 'liaison people': facilitating a research utilisation intervention in policy agencies. *Evid Policy*. 2016. <https://doi.org/10.1332/174426416X14817284217163>.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

