MAIN ARTICLE

Techniques to enhance the public policy impact of qualitative system dynamics models

G.A. (Guido) Veldhuis,* E.M. (Eefje) Smits-Clijsen and R.P.M. (Rob) van Waas

Abstract

This article demonstrates techniques to enhance the public policy impact of qualitative system dynamics models. We focus on the effective use of a large causal loop diagram (CLD) to explore a multifaceted problem situation. We discuss the conditions that can lead to developing a large CLD, the challenges this presents, and techniques that can be used to overcome them. Several techniques are discussed related to an online group model-building (GMB) process, the use of quantitative data, visual model analyses using a software tool and reporting. The techniques are demonstrated using an impactful case study on the social impact of the COVID-19 pandemic. We reflect on the efficacy of the approach through the lens of systems thinking and conclude that the techniques made a positive contribution to all aspects of systems thinking. Several avenues for future work are discussed.

Syst. Dyn. Rev. (2023)

Introduction

Since the end of 2019, the world has been grappling with the COVID-19 virus. As the pandemic intensified, the effects of the virus and government measures on society were becoming evident, with a multitude of effects still being anticipated. System dynamicists joined the effort to develop understanding of the dynamics of the pandemic (Ghaffarzadegan and Rahmandad, 2020; Struben, 2020). Their work provided new insights in the epidemiological dynamics involved. Few would disagree that the impact of the COVID-19 crisis is rife with complex dynamic behaviour beyond that of the virus. The pandemic was affecting health care, education, the economy, public safety, and many more social domains. Healthcare avoidance, social isolation, and business closures are but a few of the social effects of the crisis. The multifaceted nature of the crisis meant that identifying and understanding the wide variety of effects and their interrelations posed an unprecedented challenge for policymakers. The need for a systems perspective was more clear and acute than ever.

Systems thinking involves investigating the dynamic behaviour of a system, exploring the origin of that behaviour within the system, seeing the big picture, and diving deep into the operations and circular causation that shape the problem behaviour (Richmond, 1997). The system dynamics (SD) method seemed uniquely suited to support policymakers to develop one integrated perspective on the effects of COVID-19. However, the multifaceted nature of the crisis presented a challenge from a modelling perspective. A best practise in modelling is

TNO-Netherlands Organization for Applied Scientific Research, The Hague, The Netherlands

* Correspondence to: G.A. (Guido) Veldhuis, TNO-Netherlands Organization for Applied Scientific Research, Oude Waalsdorperweg 63, 2597 AK The Hague, The Netherlands. E-mail: guido.veldhuis@tno.nl

[Qualitative Modeling and Data, Systems Thinking, Public Policy, Causal Loop Diagrams, Group Model Building, COVID-19].

Accepted by Krystyna Stave, Received 31 August 2022; Revised 22 June 2023 and 15 November 2023; Accepted 27 November 2023

System Dynamics Review
System Dynamics Review
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/sdr.1758

starting from a clear and precise problem definition that serves as a "logical knife" to cut out all that is not relevant for solving the problem (Sterman, 2000 p. 89). In the case of the social impact of COVID-19, selecting one or a few problems as a focus would not do justice to the broad, tightly coupled, and rapidly evolving set of problematic social effects. The alternative was to explore many facets of the crises, which would inevitably lead to a broad scope and a large model. Supporting policymaking is at the core of SD, but as Ghaffarzadegan *et al.* (2011) observes: "Despite the high applicability to public policy problems, System Dynamics is currently not utilized to its full potential in government policy making" (2011, p. X). Where Ghaffarzadegan *et al.* (2011), Forrester (2007) and others have argued for the value of small models to influence public policy, we will discuss how a large qualitative model can benefit policymakers.

This article discusses an approach by which a large causal loop diagram (CLD) can offer value for the public policy arena when faced with a multifaceted problem. We will do so based on our experience supporting the Dutch government in exploring the social impact of the COVID-19 pandemic. The broad problem scope led us to use a large CLD. The use of this model type is perhaps met with some scepticism in the SD field. Therefore, we illustrate several techniques that can enhance the impact of large CLDs in the public policy arena. The techniques are aimed at the model-building process, the use of quantitative data, the use of visual analysis software, and effective visualization and documentation. Throughout we reflect on the ability of the approach to support systems thinking following Richmond's (1997) popular and comprehensive set of skills (Stave and Hopper, 2007). We then describe the contribution of the study to Dutch COVID-19 policy. In the discussion, we reflect on the approach and discuss avenues for further refinement.

Systems thinking

In the following sections, we will describe the method, process, and specific techniques used to conduct the project. While doing so we will reflect on how these contribute to supporting systems thinking in the public policy arena. We will structure our reflection based on seven critical thinking skills proposed by Richmond (1993, 1997):

- *Dynamic thinking*: "Dynamic thinking is the ability to see and deduce behaviour patterns rather than focusing on, and seeking to predict, events" (Richmond, 1993 p. 122).
- *System-as-cause thinking*: System-as-cause thinking aims to formulate a dynamic hypothesis about the problem behaviour originating from within the system as opposed to from an external force (Richmond, 1997).
- *Forest thinking*: Forest thinking is about seeing the big picture (the forest) as opposed to focusing on details (the trees). Forest thinking inspires focusing on the major themes, trends, and factors within a system and avoids being distracted by too much detail (Richmond, 1997). Forest thinking helps to establish the proper level of aggregation for a model.
- *Operational thinking*: "Operational thinking tries to get at causality—how is behavior actually generated?" (Richmond, 1997 p. 3).

DOI: 10.1002/sdr

- *Closed-loop thinking*: Closed-loop-thinking skills are aimed at seeing "causality as an ongoing process, rather than a one-time event" (Richmond, 1997 p. 3). Closed-loop thinking involves organising causal hypotheses in circular "loops" as opposed to linear chains of cause-and-effect.
- *Quantitative thinking*: Quantitative thinking aims at being numerically precise about assumptions, even if variables might be hard to measure. "You can always quantify, though you can't always measure" (Richmond, 1997 p. 3).
- *Scientific thinking*: Scientific thinking is the act of actively gathering and processing information to disprove a hypothesis. "The current prevailing wisdom is always regarded as merely an 'entertainable hypothesis'" (Richmond, 1997 p. 4).

The social impact of the COVID-19 crisis in the Netherlands

Project initiation

We were contacted in the fall of 2020 by the Dutch Ministry of Social Affairs and Employment on behalf of a group of ministries and local governments. A committee report had concluded that vulnerable societal groups were experiencing a multitude of adverse effects (VNG, 2020). In response, the government was seeking ways to further refine its reaction to and recovery from the COVID-19 crises. Having worked with us before, the client was aware of the potential of SD and group model building (GMB). In a series of meetings, we started defining and planning a project that would involve a large number of stakeholders confronted with many related problems across vastly different policy domains.

In addition to this, within every domain the associated agencies, public bodies and institutes were collecting and analyzing their own data. An integrated picture was absent. Most studies conducted at that point focused on a single domain, such as the effects of increased time at home on child abuse (Vermeulen *et al.*, 2020) or effects of school closure on learning loss (Engzell *et al.*, 2021), while others investigated which factors would protect people from the mental impact of COVID-19 (Snel *et al.*, 2021). Data were gathered by national institutes on wellbeing, support for government policies, trust in institutions, and other factors (RIVM, 2021; De Klerk *et al.*, 2020), while some municipalities conducted their own research (Gemeente Amsterdam, 2020).

The client and research team identified a need for governmental organizations to exchange, synthesize, and interpret information that was scattered across layers of government and functional stovepipes. We were requested to aid in this effort by creating a shared problem perspective and information position amongst a myriad of governmental organizations. The project aimed to enable analysis, interpretation, and anticipation of societal effects observed and still to come. Since there were many different interests at play, we suggested a common overarching goal to the actors involved: to prevent the emergence or aggravation of social problems in Dutch households as a result of the COVID-19 crisis.

Epidemiological dynamics and policy advise were placed out of scope because they were addressed by a separate part of government. The work was done assuming that COVID had spread, and that measures had been taken to control

DOI: 10.1002/sdr

that spread without explicitly representing the spread or the way those measures responded to it. This enabled the current study to focus on the social effects of the pandemic and their interaction.

Approach

The purposes qualitative and quantitative models can serve and the constraints under which they are used differ. In this study, we chose to develop a qualitative model. A qualitative model describes system structure but does not possess the ability to computationally simulate the behaviour of that system. Qualitative models are particularly useful to explore a broad problem, describe system structure, summarizes large amounts of information, support dialogue amongst stakeholders, facilitate learning, and set a policy agenda (Coyle, 2000, 2001; Vennix, 1996; Wolstenholme, 1982, 1999). Developing and using a qualitative SD model can assist actors in the policy arena to take a step away from business as usual, the use of one-way causal chains and relying on intuition and speculation (Lane, 2008; Richardson, 1999; Wolstenholme, 1999). A qualitative SD model can be constructed using a stock-and-flow diagram or a CLD. In this case a CLD was used. Given the project goals, the CLDs ability to provide a high-level overview of the feedback structure was preferred over a more detailed description of the locations and operations of stocks and flows. CLDs emphasize the presence of feedback structure and were considered a more accessible diagramming technique for the actors involved (Lane, 2008).

Amongst several limitations, the absence of simulation to test assumptions has been pointed out as a major limitation of qualitative models (Homer and Oliva, 2001; Lane, 2008; Morecroft, 1982; Richardson, 1986, 1999; Schaffernicht, 2010; Sterman, 2000). In this case, simulation was both out of reach and not a necessary condition for success given the multifaceted nature of the problem, project constraints, and project goals. Even so, in an unprecedented pandemic, special care was required to separate facts from fears using both scientific and quantitative thinking. Therefore, we used process techniques and data to support the model-development process and result with the goal to move as far away from "intuition and speculation" as possible (Wolstenholme, 1999 p. 424).

The model was developed following a GMB process. GMB is an effective process technique to facilitate a group to exchange information, carefully reflect on their assumptions, engage in systems thinking, and develop a shared problem perspective (Rouwette *et al.*, 2002; Scott *et al.*, 2016). The GMB process revolved around a series of workshops in which our team used a selection of the techniques available in the field (Richardson and Andersen, 1997; Vennix, 1996). Four meetings of 3 hours were organized in which between 27 and 38 people participated (see Table 1). The participants were mid- to senior-level civil servants and researchers who were tasked with monitoring the crisis and providing policy advice. A broad representation of different interest is essential when applying SD to inform policy, as the modelling process itself will to some extent becomes a part of the political process (Malbon and Parkhurst, 2022). Therefore, participants from a wide variety of ministries, national institutes, universities, and municipalities were invited. A portion of the sessions was used by the client to introduce

DOI: 10.1002/sdr

Table 1. Overview of the group model-building process

Objective	Plenary	Breakout rooms (BR)
Session 1: February 5, 2021.		
Preparation: Project objective and approach, initial literature scan of social effects.		
• Identify the possible and/or observed effects of the pandemic and discuss the dynamics of these effects.	• Introduction by client • Project introduction • Individual assignment: Enter effects in XLEAP • BR groups: Present most important, uncertain, and interesting effect. • Plenary discussion on dynamics	• One BR per domain; participants allocated based on expertise • BR assignment: Review effects, enter potential causes for each effect in the domain in XLEAP
Session 2: March 4, 2021 (focus on family dynamics and children)		
Session 3: March 11, 2021 (focus on adults)		
Preparation: Extended literature scan; initial model based on literature and session 1.		
Model development of variables related to either adults or families and children.	• Introduction by client	• Three BRs which discussed two domains each; participants allocated based on expertise.
	• Model presentation	
	• Poll and discussion regarding the vulnerability of societal groups/children	• BR assignment: Improve and extend the model. Facilitate model building
	• BR groups: Provide back brief on most important insights	• Which effects are observed or possible (i.e., feared)? Which effects are perceived to be temporary, and which are expected to be enduring? How do variables interact to create the observed or possible effects? How do effects differ for certain societal groups? Which variables reinforce each other, why, and how? Which variables dampen other effects?
	• Poll and discussion: Are the most important variables included in the model?	• In addition, the BR was asked to specifically check a selection of relations in the model.
Session 4: April 1, 2021		
Preparation: Final concept model, definition of 11 problematic “focus areas”		
• Consensus on model development.	• Introduction by client	Three BRs focused on four focus areas each
• Consensus on focus areas including associated risk factors.	• Model presentation	• Is the focus area well defined (effects, causes, consequences, feedback loops)?
• Initial policy exploration and way ahead	• Discussion on model	• Which personal risk factors are relevant for this focus area?
	• Presentation of focus areas • Discussion on way ahead	• What should we do (more) to tackle this problem?

DOI: 10.1002/sdr

and discuss the broader collaboration amongst the actors involved, such as the creation of a data dashboard (CBS, 2021).

COVID-19 forced us to conduct all GMB sessions online. Online GMB is a recent but promising development (Zimmermann *et al.*, 2021). We used a standard video conference service which enabled people to join the sessions from the safety of their home. To facilitate in-depth discussion, we made use of breakout rooms (BR) in every session. The service enabled us to instantly place people in BRs or recall them to the plenary room. This made switching between different program segments very efficient. The research team only consisted of three people, meaning only one team member was available per BR. This meant that the researcher had to combine several roles that would normally be shared amongst team members, most importantly the role of modeller and facilitator (Richardson and Andersen, 1995). Although not ideal, the online format made it easier to combine these roles as the team members were seated comfortably and did not have to devote attention to activities such as moving around the room. In addition, participants were assigned tasks, such as entering data or summarizing the discussion.

In the following section, we discuss the execution of a GMB process aimed at exploring a multifaceted policy problem in more detail (see Table 1). We will discuss the process used and where we deviated from standard practise. After this section, several specific techniques are introduced that were used during and outside of the GMB sessions.

The modelling process

Initiating an exploratory modelling process

Gathering reference mode data is considered a best-practice in SD modelling (Martinez-Moyano and Richardson, 2013). We anticipated that we would have to gather, organise, and verify a large number of social effects making a structured approach a necessity (Farr *et al.*, 2022). The initial set of reference modes was gathered using a literature scan of academic and grey literature. We gathered, labelled, and analyzed social effects that surfaced from the literature scan and GMB sessions in a table. The table was used before and during model building to identify model domains, identify and aggregate variables, and triangulate sources. An intermediate step between data collection and modelling prevents a large model from quickly growing beyond the control of the modellers in addition to the benefits laid out by Farr *et al.* (2022), such as separating source data from inferences by the modelling team. We predominantly searched for sources using Dutch data with google search and google scholar. We used snowballing to extend the analysis via references and citations. Due to the large number of effects and limited availability of longitudinal data, we expressed reference modes in the table in simple terms: an increase or decrease in a variable and whether this effect was observed or was assumed to occur (scientific thinking). Each effect had the following attributes:

DOI: 10.1002/sdr

- *Domain*: Eventually, seven domains were identified: physical health, mental health, family and development, social cohesion and participation, wellbeing and resilience, public safety, work and income.
- *Effect*: A simple description of change over time—for example, decreasing premature births, postponement of surgeries, decrease of cancer diagnoses.
- *Cause(s)*: A brief description if causes for the effect had been suggested or established by other sources.
- *Consequence(s)*: A brief description if consequences of the effect had been suggested or established by other sources.
- *Affected group*: A note if the effect or source was related to a specific societal group.
- *Appreciation*: A note if the effect is normatively positive, negative, or neutral.
- *Observed/assumption*: Sources that use quantitative or qualitative empirical evidence were labelled as "observed" while source that used expert opinion were labelled as assumption.
- *Sources*: Literature references, and later, references to statements made in GMB sessions.

Peer-reviewed publications were scarce, but other reputable sources were available. We found that quite some research was already published in the first year of the pandemic in the form of research reports by institutes and government sources. Some examples have been provided in the project-initiation section. Our preliminary analysis was based on 43 sources. During the project the initial literature scan was extended to 122 sources of which 95 were used to support the final result. Each of these sources contained numerous data points. Major social challenges, like the COVID-19 crisis, present policymakers with the challenge of having to identify the major themes of the situation amidst the noise of many datapoints, news headlines, political debate, and interest groups. Getting the data supports quantitative thinking. Distinguishing between observed data and assumptions encourages scientific thinking. Further organization of data is a first step towards identifying variables and supporting forest thinking to derive a bigger systemic picture. In this case, a large volume of information and perceptions were available about "the trees" but little shared understanding existed about what the "forest" looked like. Few of the sources separated between causes and effects, neither empirically nor with theory. This signals that there was still relatively little insight available in the operations and loops driving the system behaviour. We aimed to close this gap by using a combination of data and expert knowledge elicited during GMB. The table of effects formed the starting point for the first GMB session. The use of data in the sessions will be discussed in more detail in the section entitled "Quantitative data and exploratory CLDs".

First session: Exploring a multitude of reference modes

The goal of the first session was to identify the observed and assumed effects of the pandemic, discuss their dynamics, and explore their causes. In a frequently used GMB exercise, participants are asked to sketch and discuss behaviour over time graphs (Andersen & Richardson, 1997). Unfortunately, sketching and discussing graphs was not an option due to the online setting, time availability, and

DOI: 10.1002/sdr

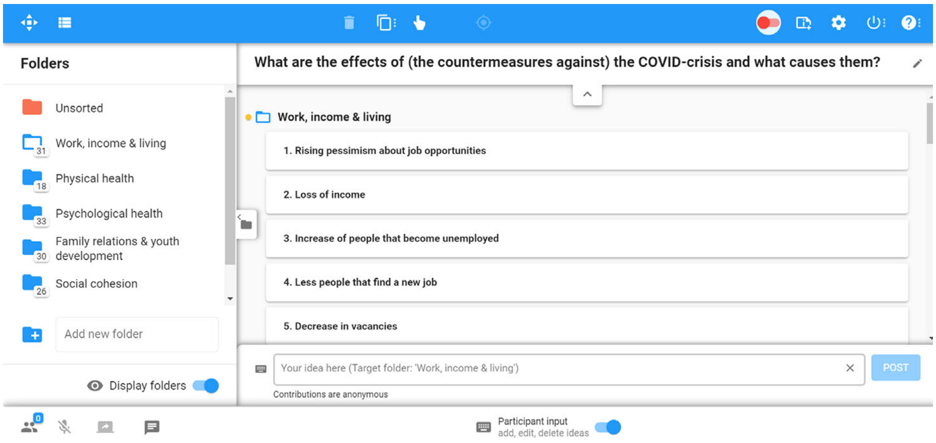


Fig. 1. Print screen of the X-leap environment. On the left side is a navigation pane with folders for each domain. On the right side, the effects in the current folder are displayed. Participants can add new effects or respond to existing effects.

anticipated large number of effects. We did however wish to provide the space for all participants to contribute their observations about the effects of COVID-19 and encourage them to engage in dynamic thinking. To achieve this, we developed an alternative approach using the online collaboration tool XLeap[i] (see Fig. 1).

The effects identified by the literature scan were entered in the XLeap environment and organized in folders according to domain. Participants were asked to supplement and review the list. Each participant was able to add an effect or add information to an effect. Using the existing table of effects had several benefits: it saved time since many effects were already entered, and participants had the opportunity to reflect on and respond to the results of the literature scan. The exercise was productive and led to more than 60 additional effects being identified. Having participants enter their own thoughts comes with the benefit of engaging them in the process and creating a sense of ownership (Andersen and Richardson, 1997; Vennix, 1996). Actively formulating and categorizing effects helps to establish both depth and breadth of the model and is a first step in engaging the group in systems thinking. To encourage participants to think dynamically, we asked to enter their expectations about when the effect would peak (it already peaked; is currently peaking [February 2021]; will peak in summer of 2021; will peak end of 2021; or will peak in or after 2022). Applying dynamic thinking also encourages forest thinking as people are encouraged to think beyond individual events.

After a break, the participants were divided across BRs based on their expertise; each room focused on two domains. We asked participants to review the effects and discuss causes. The participants were asked to enter the insights from their discussion directly into XLeap. One participant from each BR was asked to brief the full group on the most interesting, important, and uncertain effects discussed. In addition to dynamic thinking, this exercise encourages forest thinking as participants review categories and select variables and their dynamic behaviour, and asking participants to formulate what causes an effect that encourages

[i] www.xleap.net.

DOI: 10.1002/sdr

operational thinking. This is illustrated by the following quote in which a participant describes system behaviour while also trying to determine a relevant level of aggregation as well as the operations involved: “We looked at wellbeing, there are allot of factors listed that are strongly correlated, we should cluster those: Increase in fear, stress, gloom, a decline in well-being … We believe the cause of this to be related to the uncertainty that’s forming ⋯ also the duration of the crisis, but also financial uncertainty ⋯ those are all factor that cause ⋯ but up until some level, everyone is a bit gloomy at the moment, but when does that become problematic?” Participants shared many similar insightful thoughts in which they used one or several systems-thinking skills.

Second and third session: Model development

In between session 1 and 2, a preliminary CLD was developed by the research team based on the literature scan and results from session 1. Modelling was done in the MARVEL software, an in-house modelling tool (Veldhuis *et al.*, 2015). Starting from a preliminary or “seed” model saved valuable in-session time, stimulated the start of the discussion, and enabled us to leverage the information gathered so far (Hovmand *et al.*, 2013; Vennix, 1996). Starting from a preliminary model can pose the risk that participants do not develop a sense of ownership of the model. After carefully introducing the model and referring to the results of session 1, we noticed that participants quickly engaged with the model. It is also important here to note that this GMB exercise did not intend to capture the meaning the individual participants ascribe to the pandemic but to create a shared perspective on the social impact by triangulating their expertise and other sources. Put differently, GMB is used here predominantly in a positivist mode, not a constructivist mode (see de Gooyert *et al.*, 2019, for a discussion of the use of qualitative GMB in a constructivist mode). Session 2 was focused on adults, and session 3 was focused on children and families because various policies specifically affected children and families, like school closures, and caused concerns for their safety and development. Both sessions had a similar set-up.

During both sessions we challenged participants to reflect on the model. This took place in three BRs that each addressed two domains. Attendees were assigned to BRs based on their expertise. The facilitator shared their screen so that participants were able to see the model and changes in real time. We used several questions to kick-off the discussion and trigger dynamic, operational, and closed-loop thinking—for example, which effects are perceived to be temporary due to the measures against COVID-19, and which are expected to be enduring? (See Table 1). This question elicits effects but requires more cognitive effort than simply producing a “laundry-list” of effects. It challenges participants to think about effects but also about the underlying operations that cause an effect to be temporary or enduring. Participants often described one or several personal risk factors (such as having a disability) which they believed would moderate the effect. The questions (see Table 1) formed a starting point for a facilitated discussion on both existing and new variables and relations. During these discussions, the facilitator/modeller assisted participants to voice both causes and consequences of the variables being discussed (see Vennix, 1996, p. 120). Addition and changes to the model were made on the basis of consensus.

DOI: 10.1002/sdr

In addition to allowing the group to form its own thought process, we also wanted to use their scientific-thinking ability to test certain specific assumptions. Before the session, a set of variables and relations had been selected which were included in the model but for which the evidence was unsatisfactory. If needed, the facilitator explicitly asked participants to reflect on them and decide if they should be included.

During the discussions, we asked participants to reflect on whether their assumptions could be verified with either quantitative or qualitative data, encouraging them to apply scientific thinking as part of discussing the operations and loops in the system. When assumptions seemed to conflict with available data, we entered into a discussion on what this might mean for the assumptions or for the data (see the section entitled "Quantitative data and exploratory CLDs"). During such discussions, we use software that visually highlights causal paths and loops in the model. We will discuss this technique in more detail in the section entitled "Cutting through the visual complexity of large CLDs." After the sessions, the insights from the breakout rooms and adjusted CLDs were combined to create an updated version of the model. This was relatively straightforward as each BR focused on specific domains. Conflicting changes from input gathered in separate BRs were resolved by the modelling team. Some changes required additional expert input; this led us to interview participants in between sessions. Interviews were conducted online while the model was visible on screen, and changes could be made directly together with the participants.

Fourth session: Consensus on results

The aim of session 4 was to establish consensus on the variables, relations, and feedback loops in the model as well as the selection of focus areas. In addition, a preliminary discussion was held on the implications for policy. Based on the literature scan and discussions during the first three GMB sessions, we identified one or more *focus areas* in each domain. Each focus area described a dynamic problem that was observed or expected to be problematic for the Dutch society. Examples were "postponed healthcare and blind spots in healthcare needs," "polarization and trust in government," "pressure on wellbeing," and "resilience." The focus areas were introduced to offer a more manageable way to communicate the key findings from the study. For each focus area, several risk factors were identified and discussed that make the occurrence of effects more likely, enduring, or severe. Examples of risk factors are being self-employed, having a vulnerable health, and poor language and digital skills. The risk factors helped policymakers to translate the quite aggregated systems insights into more targeted policies.

Throughout the process qualitative and quantitative data were triangulated to build confidence in the variables and relations within the model. A few examples of this triangulation are discussed in the next section. After the fourth GMB session, the model was perceived as being representative of the problem situation by the client and other participants in the study. Some effects, like domestic violence (see the following section), remained uncertain as contradicting expert assumptions and data were available. With regards to other effects, participants

DOI: 10.1002/sdr

stressed that it “might take years for all social effects to materialize.” The report reflected such uncertainties (Veldhuis *et al.*, 2021).

After GMB sessions, the analysis was finalized and communicated. The recommendations that followed from this work were mostly exploratory, in line with the purpose of the study. Recommendations included a call for a more systematic approach to analyses and better cross-government collaboration, as well as agenda setting with regards to a set of focus areas and potential risk factors. The final report was reviewed by several participants who again reflected on the question if the evidence sufficiently supported the conclusions and if the result was useful. In their response they commented that the report “substantially deepened and provided context to existing insights,” was “of great value,” and “especially the description of feedback loops[,] is useful.” The limitation of qualitative modelling did not escape the participant who voiced their interest to pursue further quantitative thinking to: “measure the social impact,” determine “which factors have the most impact,” and establish “which intervention are most effective to cut vicious cycles.” The report became an important source for policymakers as we will describe below. But first we will discuss several techniques used throughout the process to enable the effective use of a large CLD in the public policy arena.

Techniques for effectively using exploratory CLDs

Quantitative data and exploratory CLDs

In this section, we discuss the importance of using quantitative data when developing and reporting qualitative models. To our knowledge, few qualitative SD case studies present quantitative data or describe how quantitative data were used in the modelling effort.

Data draw the attention towards the dynamics of the system. In the report, model structure is presented together with over 90 behaviour-over-time graphs and many more cross-sectional statistics (examples are provided in Figs. 2, 3, and

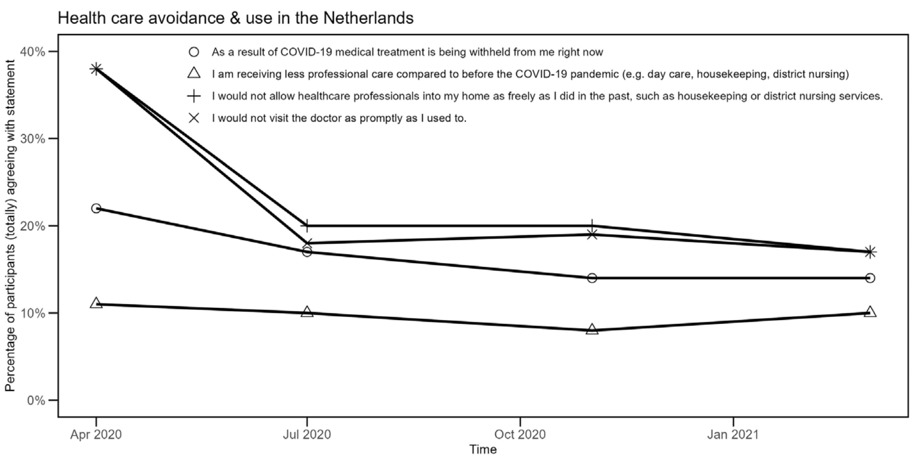


Fig. 2. Survey data on healthcare use and avoidance (Engbersen *et al.*, 2021). Lines show fraction of respondents (totally) agreeing with statements given below the graph.

DOI: 10.1002/sdr

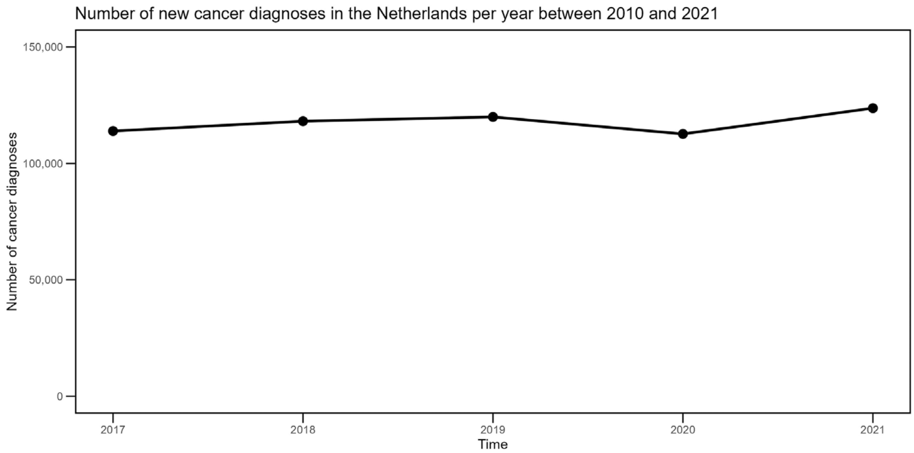


Fig. 3. Number of new cancer patients per year 2016–21 (IKNL, 2021).

5). Although the importance of behaviour-over-time graphs are often emphasized in SD, cross-sectional statistics should not be ignored by modellers. Such statistics encourage quantitative thinking and provide a perspective on the relative size of effects which are often a result of behaviour over time. For example, analysis revealed that 32% of employees or 2.4 million people worked in a sector that faced a large or very large economic decline during the crisis (UWV, 2020)—a result of a change over time that clearly had to be explored during GMB and included in the model.

Introducing data during GMB stimulates discussion and helps the participants to connect model structure to problem behaviour. Assumptions voiced during the session about system structure and behaviour can be cross-checked with data. This is a two-way street: the model can also provide more insight into the data because it contextualizes the data and positions it in relation to the system. We will explain these points in more detail with three examples: healthcare avoidance, bankruptcies, and domestic violence.

First, the data on healthcare avoidance. During the crisis measures were taken to limit the contact rate. At that time, the contagiousness of the virus was still uncertain, and the healthcare system was overburdened. Appointments were cancelled or postponed, and a part of society was reluctant to request care because of fear of contracting the virus (and spreading it to family members). These sentiments were visible in survey data (Fig. 2) (Engbersen *et al.*, 2021). A significant percentage of people indicated that they (totally) agreed with two statements of healthcare avoidance (38%–17%; 38%–17%) and/or with two statements expressing they felt that health care was less accessible to them (22%–14%; 11%–10%). Over time this effect decreased as lockdown measures were lifted and the perceived risk of the virus declined.

Although the initial effect of healthcare avoidance decreased, the follow-on effects were much more enduring. A local hospital reported 68 patients with complications because of postponed health care (Trouw, 2021). But perhaps even more worrisome is that the number of new cancer diagnoses decreased for the first time in the last 30 years (Fig. 3, IKNL, 2021). In an interesting example of operational thinking, a participant explained that this disrupts the normal

DOI: 10.1002/sdr

operations of the healthcare system which is geared towards a steady flow of patients through limited diagnostic and treatment capacity. Disruption of the flow causes a, partly hidden, backlog which might be hard to clear later while delayed diagnosis and treatment can severely impact the prospect for patients. The National Institute for Public Health and Environment estimated that approximately 34,000 to 50,000 healthy years of life had been lost because of the postponement of treatments (van Giessen *et al.*, 2020).

The data were discussed during the GMB sessions, and the research team and the participants formulated a system-as-cause explanation of the decrease in physical (regular) health care provided to patients during the crisis. Avoiding health care by patients could be caused by fear of contracting the virus or lower perceived healthcare accessibility. The avoidance of care together with delayed care causes blind spots in current and future healthcare needs since diagnoses fell below their usual rate. Together with other causes, such as overburdened healthcare professionals, this resulted in reduced and delayed health care with potentially severe health consequences. The final version of the physical health domain is shown in Figure 4.

Having data available as conversations emerge during GMB sessions creates the opportunity to triangulate statements and to further deepen the conversation. Containment measures and other effects were expected to cause large revenue

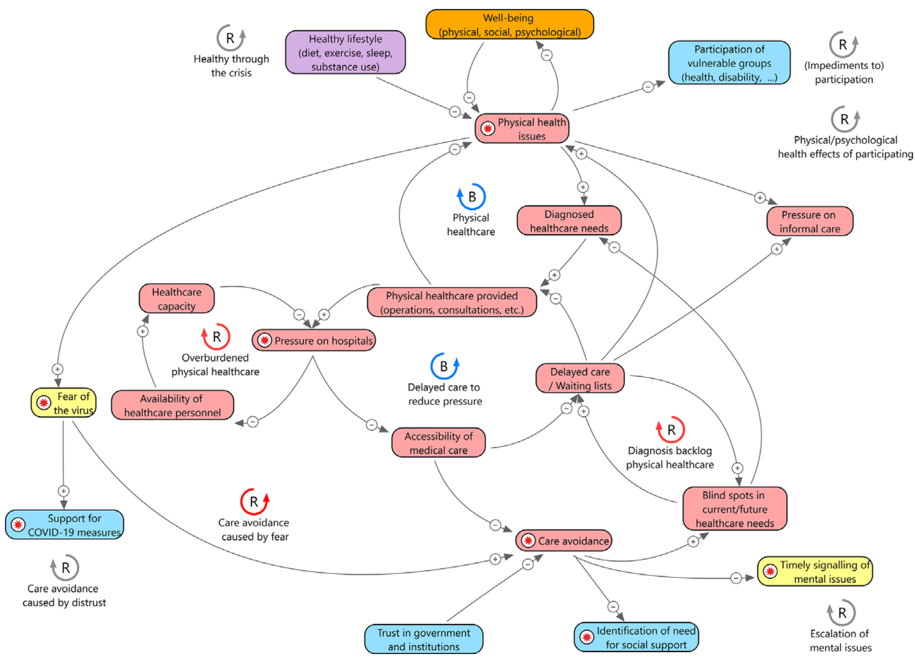


Fig. 4. The submodel of physical health. Excerpt from the full model that only includes variables in the physical health domain (red variables) and the first-degree connections to other submodels (purple, orange, yellow, and blue variables). Named feedback loops are indicated by red (reinforcing) or blue (balancing) loop symbols. The gray feedback loop symbols on the edge of the submodel indicate the presence of loops which cross the submodel boundaries and are only partially visible in the figure. Some variables have a red star that indicates that these variables are directly impacted by variables not included in the model, such as the virus and government measures

DOI: 10.1002/sdr

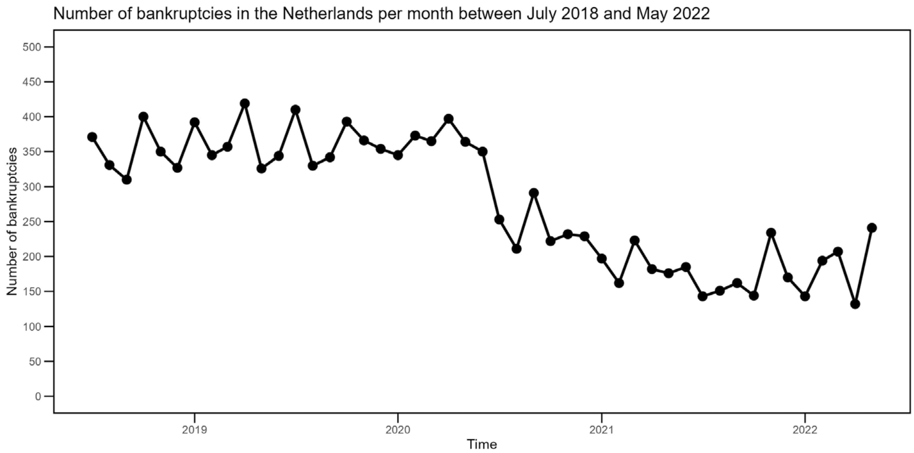


Fig. 5. Number of bankruptcies per month in the Netherlands (CBS, 2022).

losses for businesses in some economic sectors (UWV, 2020). As a results, some participants feared this would cause a dramatic increase in bankruptcies. Interestingly enough, the data gathered showed a drop in the number of bankruptcies (Fig. 5). Data were brought forward verbally in the session to deepen the conversation. In this example and other cases like this, we did not introduce the graph itself in the GMB session as this might take too much time and the main point is simple. According to this data: bankruptcies have decreased, not increased. The discussion continued and the following system-as-cause hypothesis emerged. Government-support schemes had pushed bankruptcies below a level they would normally be. In addition, additional data indicated that many entrepreneurs used their capital and personal savings to remain solvent. Both mechanisms were included in the model and supported by data in the report.

But what if expert assumptions and data seem to conflict? Many participants expected domestic violence to peak during lockdowns as tensions in families grew and social work was constrained. This dynamic was included in the model. In their reasoning for a rise in domestic violence, participants proposed an operational and closed-loop arguments. However, the data presented a mixed picture. Reported cases of domestic violence did not increase. However, the number of reports per case had increased as well as the number of reports by "informants," such as teachers (NSCR, 2021; Vermeulen *et al.*, 2020). Being able to refer to data in the session facilitated quantitative and operational thinking. A much deeper discussion emerged about how various sources used different units for reporting (per case or per report), how data was collected (neighbors who are at home are more likely to report a disturbance), which dynamics might influence the measurement (worried informants), and how the problem dynamic itself was influenced by the crisis. Our report presented the findings of the discussion and data analysis and discussed the uncertainties involved.

Cutting through the visual complexity of large CLDs

There is no doubt that large CLDs can be challenging to read and comprehend. We therefore employ a variety of techniques to draw and visually analyse the

DOI: 10.1002/sdr

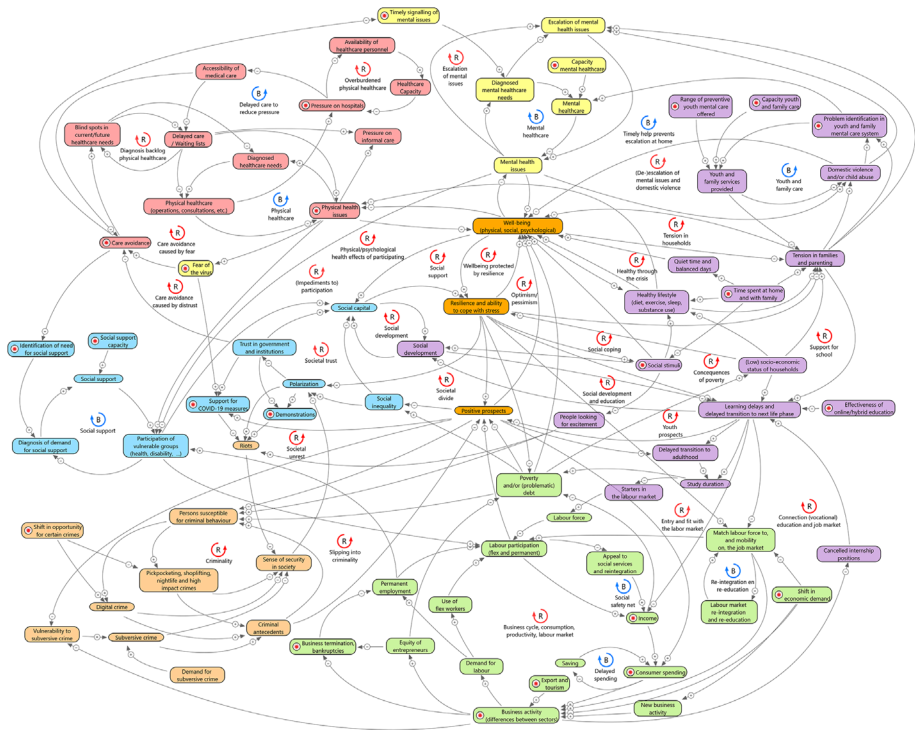


Fig. 6. The complete social impact of the COVID-19 crisis model. Variables are colored according to domain: Physical health (red), mental health (yellow), family and development (purple), social cohesion and participation (blue), wellbeing and resilience (orange), public safety (peach), and work and income (green). Some variables have a red star indicating that these variables are directly impacted by variables not included in the model, such as the virus and government measures

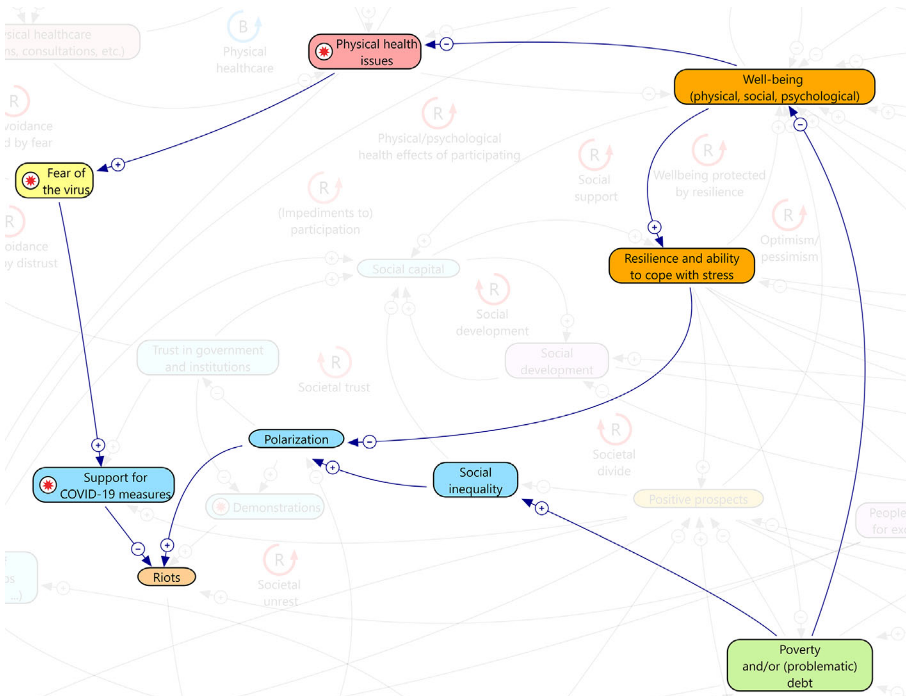


Fig. 7. An example of the path tool in action. The tool has identified and highlighted three causal paths between two variables selected by the user: "Poverty and/or (problematic) debt" and "Riots." Visible are variables from the domains work and income (green); physical health (red); mental health (yellow), wellbeing and resilience (orange); social cohesion and participation (blue), and public safety (peach)

DOI: 10.1002/sdr

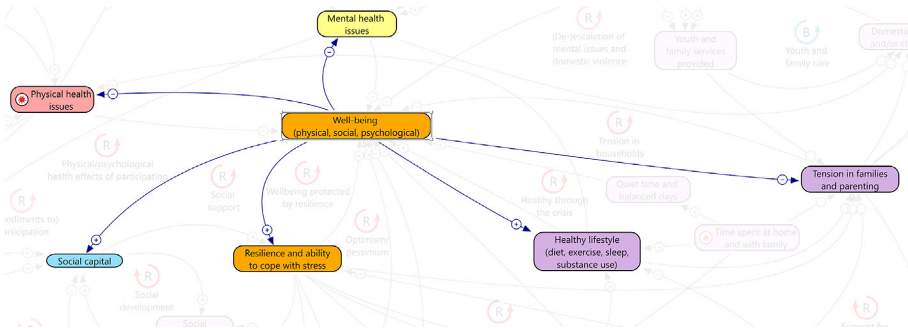


Figure 8. Radiate visualization of the first-degree outgoing relationships from the variable "wellbeing"

model. These techniques make the model more accessible, easier to analyse, aid in communicating insights, and help to focus attention during GMB sessions.

We use submodels for communication purposes (see Figs. 4 and 12). A submodel is an excerpt taken from the larger model. In our experience the less variables a model has the easier it is to organise properly and communicate. A risk of using submodels is that the audience loses track of the relations of a submodel to the larger whole. This can especially present a problem if the submodels describe functional domains, since especially the interdomain relations might contain new insights. For this reason, we include variables at the boundary of the domain in the diagram and include feedback loop symbols of loops that cross the submodel boundary.

The use of color in SD diagrams is not standard but, in our experience, has several benefits. Colors aid an audience to orient themselves within a model. In the social-impact model, variables are organized and colored according to seven domains (Fig. 6). This creates an extra layer of information by which a reader can quickly identify to what policy domain a variable belongs and by which domains it is influenced. It aids forest thinking by providing the reader with an aggregated level of information on top of the level of individual variables and relations. When using submodels (see Fig. 4), the colors are a way for the audience to position the submodel within the full model. The color-coded domains are

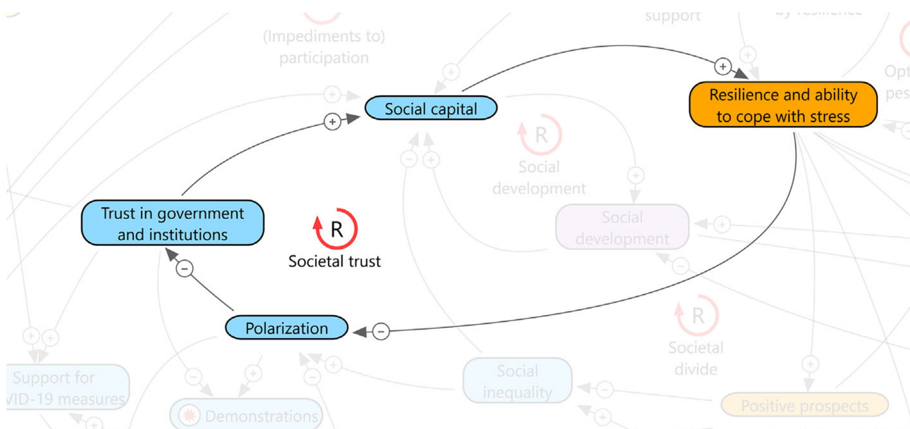


Fig. 9. The reinforcing loop social trust, highlighted using the loop tool

DOI: 10.1002/sdr

subsequently used to visually link the model to report sections and other visuals used in the project (see Fig. 13). In our case study, we observed these qualities anecdotally when participants were able to quickly switch between various visualization and trace how a domain was connected to other domains.

As diagrams grow larger, it becomes more difficult to trace the causal connections between variables. To overcome this problem, we use visual analysis tools that automatically highlight variables and relations based on a selection of variables. We use three tools: radiate, paths, and loops. The tools make a selection of variables and relations stand out by temporarily changing the color of the relations that are part of the selection and by reducing the opacity of variables and relations that are not a part of the selection. This enables users to view a specific selection of variables and relations while not losing their orientation on the larger model. We use these tools continuously during GMB sessions and when we conduct and present analyses (see Figs. 7–9).

The path highlight tool

The path tool highlights the variables and relations between two selected variables. The user can select which paths to include if multiple paths are present. We highlight paths during GMB to visualize how two specific variables are connected. Often narratives put forward by participants about the operations of the system include multiple causal steps or even separate paths. The path tool helps other participants to follow the narrative and helps the participant to check whether the causal paths in the model are a good representation. In addition, we often find that GMB participants are eager to add relations between existing variables. In such cases we use the path tool to highlight the existing causal paths between two variables. We then ask the participants if and how the suggested relations operate differently from the paths already present in the model. Quite often we find that participants have overlooked that an existing path already captures their reasoning. This prevents adding superfluous relations.

An example of three paths from poverty to riots is included (Fig. 7). This visualization was used in a discussion about the so-called curfew riots that occurred in 2021. It shows a selection of potential underlying causes of riots and allowed GMB participants to discuss effects stemming from various domains into the safety and security domain.

The radiate highlight tool

The radiate tool highlights all causes and/or consequences of a variable or a set of variables. The user can set the degree of separation to be highlighted (Fig. 8). During GMB sessions or presentations, this tool assists in quickly identifying how a single variable or group of variables is embedded in the larger system. During a session a participant downplayed the role of wellbeing and argued that actions within their sphere of influence had no relation with wellbeing, whereas actions in other domains did. Using the radiate tool, we could instantly show how wellbeing was deeply embedded in the problem structure. The quick visual analysis revealed how wellbeing was indeed connected to the domain in which the participant worked, both as an indirect cause and consequence. This

DOI: 10.1002/sdr

deepened the understanding of the participant and led to a productive discussion about the central role of wellbeing in the problem.

The loop highlight tool

The loop tool identifies and highlights a loop passing through a selected variable and indicates whether this is a balancing or reinforcing loop (Fig. 9). The shortest loop is presented first after which the user can scroll through additional loops. The tool helps the modeller and audience to identify which loops are presented in the model. It greatly aids closed-loop thinking by visually showing how a variable is part of the feedback structure of the model. Like the other tools, it can be used to focus attention on a single piece of structure for discussion or presentation purposes. It is particularly helpful in large diagrams which can be difficult to organise according to model layout best practises, obfuscating the location of loops. In GMB, it is common practise to draw the attention of participants to feedback loops that emerge during the conversation (Vennix, 1996). At such moments we use the tool. We also use it when the dynamics of a specific variable are being discussed to quickly search the model for loops involving that variable and to nudge participants to reflect on the feedback mechanisms and the role they play in the behaviour. During the GMB sessions, the participants identified resilience as one of the central variables that mediated the impact of the crisis on the lives of people. The example in Figure 9 illustrates one of the feedback loops that emerged. It shows a hypothesis for the worrisome downward spiral in social trust that occurred in the later stage of the pandemic. The visualization of feedback loops also has value in documenting the model which we will elaborate on in the next section.

Documenting the model and its insights

Documentation of the model and the findings of the study is critical in transferring insights. A small group of stakeholders is involved in the GMB sessions,

Fig. 10. A high-level illustration of the seven domains of the social impact of the COVID-19 crisis model.

DOI: 10.1002/sdr

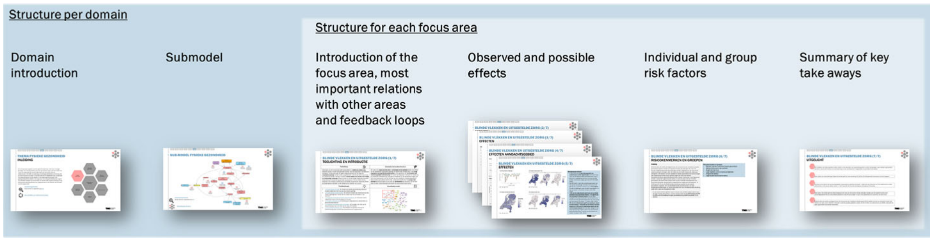


Fig. 11. Illustration of the structure in which the content of analyses was presented in the main body of the report.

while a bigger audience will depend on the documentation to gain insights from the study. A diverse group of policymakers with different backgrounds and responsibilities use the report. A tremendous amount of information is needed to support and explain a large model. At the same time, policymakers with no prior knowledge of SD will have little time to read the document and will have an appetite for to-the-point conclusions. The challenge is to organise all information in a way which is effective and efficient in communicating the findings to the audience while not losing the unique systems perspective of an SD study. The general structure of the report followed standard practise with an introduction, method, body, discussion, conclusion, and recommendations sections. In this section, several specific techniques related to the organization of the main results will be discussed (see Fig. 11).

Thematic organization and a clickable document enabled the audience to navigate swiftly between specific segments and overall findings. We use an

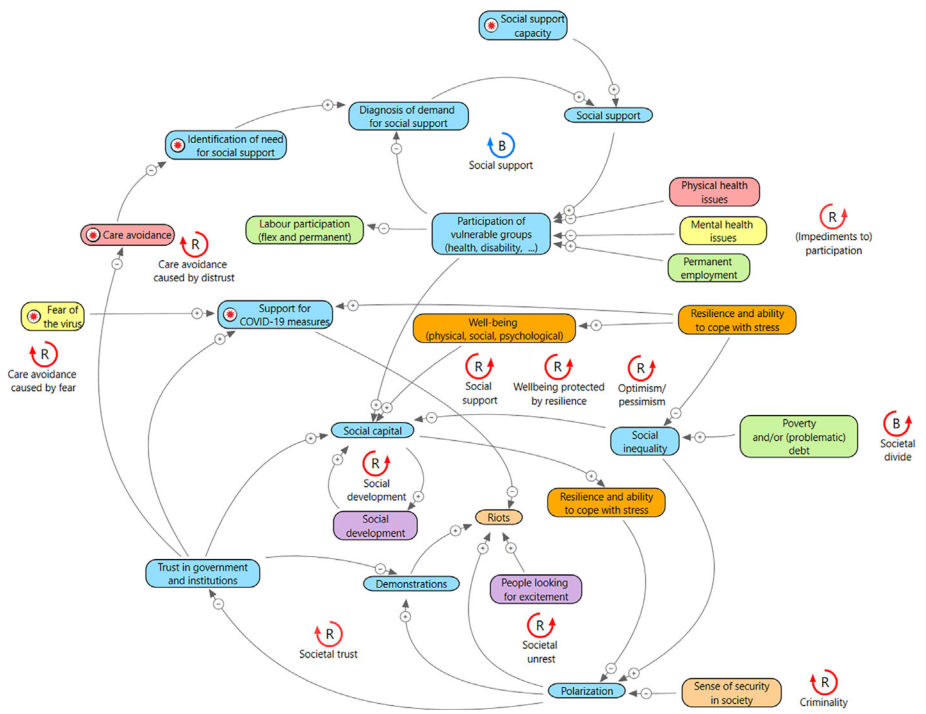
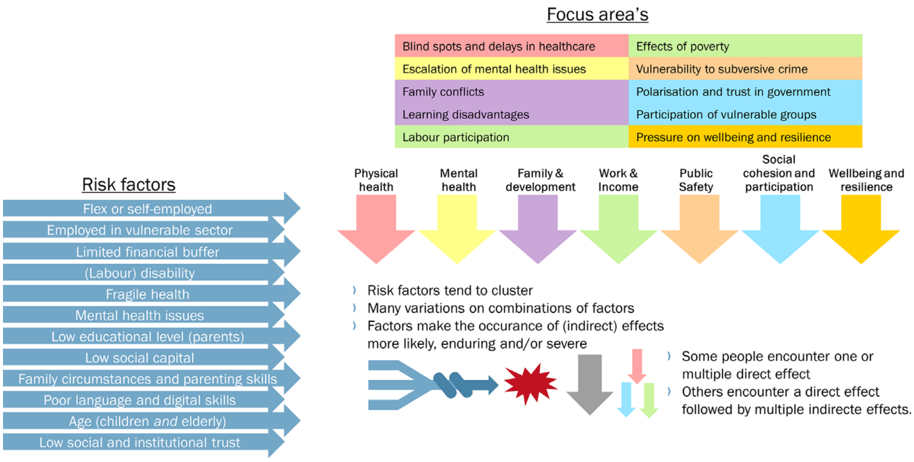


Fig. 12. The submodel of the domain social cohesion and participation. Excerpt from the full model that only includes variables in the social cohesion and participation domain (blue variables) and the first-degree connections to other submodels (purple, orange, yellow, green, red, and peach variables). Named feedback loops are indicated by red (reinforcing) or blue (balancing) loop symbols. The gray feedback-loop symbols on the edge of the submodel indicate the presence of loops which cross the submodel boundaries and are only partially visible in the figure

DOI: 10.1002/sdr

Fig. 13. A schematic summary of the main findings: seven domains, 10 focus areas for policy, and 12 risk factors that increase the vulnerability to experience effects. Interactions within and between these elements are briefly explained at the center of the figure.



additional layer of description on top of the model: domains and focus areas. Simple visualizations make the content of the model clear at a glance and serves as an important anchor point throughout the whole report (see Figs. 10 and 13).

Traditionally, SD favors model publications that display "a compact model that shows how the difficulty is being caused" (Forrester, 2007 p. 365). Reading a large model diagram can be a daunting task even for the trained eye. Bures proposes that due to "… cognitive limits, only model fragments are taken into consideration and the value of … the causal loop diagram diminishes" (2017, p. 4). Large models require gradual clarification to be understood. In the report each domain is introduced with a description and a separate submodel (Fig. 12). Each submodel is an excerpt from the complete model, but its structure is rearranged according to standard practises. This is followed by the focus area(s) of a domain, in which its operations and dynamics are discussed in several sections. Each focus area has an introduction which summarizes key dynamics in the focus area, an explanation of the interaction with other domains, an overview of the feedback loops that influence the focus area, and the location within the full model (Fig. 6). The observed and possible effects section describes the model structure, supporting information, and the dynamic behaviour of the variables. This assists the reader in relating model structure to problem behaviour and supports dynamic thinking. It also provides the information needed for quantitative thinking—for example, on the prevalence of effects in the population.

Thirty-five feedback loops in the model are numbered and have a descriptive name (see Fig. 9). The feedback loops are used in the introduction of every focus area. For an untrained audience, each feedback loop requires explanation. Therefore, each feedback loop is displayed and discussed separately in the appendix of the research report (Veldhuis *et al.*, 2021).

The amount of information communicated in an exploratory study can be overwhelming. Therefore, four to five key takeaways are shared at the end of every focus area. Key takeaways refer to the relevant feedback loops and include a link to the loop description in the appendix. By doing so, the short fragments remain linked to the larger model, and the reader is encouraged to maintain a closed-loop

DOI: 10.1002/sdr

and system-as-cause perspective. In the discussion section of the report, the high-level takeaways are summarized in one overview (Fig. 13). This overview summarizes the key points and makes them easily accessible for the audience.

Project impact

The report was offered to top officials and was discussed with policymakers from different ministries and municipalities responsible for policy across the seven domains included in the study. The report was also made publicly available. A senior official stated that the report was "a huge help in seeing relations" and "offer[ed] insight and support in developing integral policy." The client stated that the report and model became a frequent source of reference in policy discussions within the ministries. The study was used extensively by policymakers to produce a policy document as part of the governments "Recovery response program COVID-19 crisis" (Directorate-general Society and COVID-19 [DGSC-19], 2021; Dutch Ministry of Finance, 2021). The policy document summarizes the conclusions from the study and the key takeaways for each domain and risk factor. These findings are combined with other studies by policymakers to derive recommendations. The policy recommendations were brought to the attention of parliament in a joint letter by ministers of Health, Welfare and Sport and of Social Affairs and Employment (Kuipers and van Gennip, 2022). The recommendations were aimed at protecting vulnerable groups in society and mitigating the damage caused or aggravated by the crisis.

Our study contributed to both deepening and broadening the policy debate on the social impact of COVID-19. Before the study, policy recommendations to government primarily discussed the emergence of new vulnerable groups (e.g., elderly, children) with only brief analysis of the effects and underlying dynamics involved (VNG., 2020). Following the study presented here, a new set of policy documents contained a more detailed description of the social effects of COVID-19 and a deeper appreciation of the underlying dynamics (Directorate-general Society and COVID-19 [DGSC-19], 2021). In addition, the results of the study were used to select which variables to include in a government dashboard to monitor the social impact of the crisis in the years to come (CBS, 2021). In addition, a follow-on study on resilience was commissioned. Two years after the publication of the original study, it was cited in a letter of the cabinet to parliament to explain its policies with regards to minimizing the effects of the pandemic (Kuipers *et al.*, 2023). The letter briefly reiterated the conclusions of the study with regards to the dynamic nature of the social effects of COVID-19, their complex interactions, and the role of resilience and underlying risk factors in shaping this behaviour.

Discussion

In this article, we have provided a practitioner's perspective on the use of a large qualitative model to support policymakers who face a multifaceted problem. We have introduced techniques related to a GMB process, the use of quantitative data,

DOI: 10.1002/sdr

visual analysis, and reporting. With the use of a CLD and GMB, these techniques can be used by modellers and others to explore the problem and apply seven critical-systems-thinking skills. Examples from a study on the social impact of the COVID-19 pandemic were provided to support this argument. The approach described was able to support all seven critical-thinking skills to a degree. Supporting the skills of quantitative thinking and scientific thinking was restricted to careful triangulation of both qualitative and quantitative data, which can be achieved through a combination of data gathering and GMB. Adoption of the results in several policy documents provides confidence in the usefulness of the study to advance the Dutch response to the COVID-19 pandemic (Directorate-general Society and COVID-19 [DGSC-19], 2021; Kuipers and van Gennip, 2022; Kuipers *et al*., 2023). In many cases, it is advisable to strive for a small model (Forrester, 2007; Ghaffarzadegan *et al*., 2011). But, under some conditions, the needs of a client cannot be met with a small model, while a large qualitative model can provide benefits. In this article, we have illustrated that large qualitative models can be used to support policymakers and have presented a set of techniques to achieve this. Like other authors, we found online GMB to be effective (Zimmermann *et al*., 2021).

Several authors have raised the question of what purpose qualitative models could serve, especially the sole use of CLDs (Lane, 2008; Richardson, 1999). In this article, we have described one type of application: exploratory modelling of multifaceted problem situations to support systems thinking in the public policy arena. Qualitative SD models are not an alternative to quantitative SD models but have their own place within the field. A focus on what qualitative models cannot do distracts us from what they can do. The variety of purposes that qualitative models can serve is far more diverse than the example provided in this article. For instance, de Gooyert *et al*. (2019) describe the use of qualitative SD in a constructivist approach, Black (2013) describes the role of diagrams as boundary objects, and Niks *et al*. (2022) use CLDs to integrate a body of academic literature.

Therefore, one standard for the quality of a model, as suggested by Homer (2014), does not do justice to the diversity of applications in the field. However, vagueness about which techniques and criteria to use is clearly an undesirable situation, and a set of guidelines for qualitative modelling has been called for before (Lane, 2008; Richardson, 1999). A framework is needed that distinguishes the different modes in which SD is used. From that starting point possible differences in techniques and quality criteria used could be more meaningfully discussed. Previous attempts at such a framework seem to not have taken root (for example, see Lane, 1995). We have provided several techniques we use to assist our clients. Which other techniques and procedures might be used to improve the effectiveness of using qualitative models? Several authors have introduced techniques from qualitative research (see Desthieux *et al*., 2010; Eker and Zimmermann, 2016; Kim and Andersen, 2012; Luna-Reyes and Andersen, 2003; Turner *et al*., 2013). What other improvements might there be possible in the development and communication of qualitative models? For applications such as our case study, we have stressed the importance of using data in combination with diagrams. But which techniques might best be used to present and discuss data alongside a diagram?

The visualization of SD diagrams has been guided by a strict set of guidelines for decades. The use of color in this article breaks one of those convention, but in

DOI: 10.1002/sdr

our experience, comes with several benefits. In addition, we find that being able to qualitatively interact with a model diagram by visually highlighting relations, paths, and loops benefits our own ability to analyse models as well as that of participants. Diagrams are invaluable in any SD application, but their form and use seem stagnant. What other innovations might be possible to enhance the ability of diagrams to communicate insight and support systems thinking?

At present we can only offer a case study as evidence as so often is the case in the SD field. Only a limited set of studies has operationalized and measured changes in systems thinking (see Doyle, 1997; Maani and Maharaj, 2004; Plate, 2010, also see Stave and Hopper, 2007). Further operationalization and measurement of systems thinking and the transfer of systemic insight from models and the modelling process to users is needed. Such research can test the rich variety of ways in which policymaking is supported by SD and extend our toolkit.

Acknowledgement

We would like to express our sincere gratitude to Dagmar Sandee and Bas Torenvliet from the Dutch Ministry of Social Affairs and Employment and all who participated in the GMB sessions and interviews. We thank our colleague Bas Keijser for his feedback during the writing process. Lastly, we thank the guest editors and two anonymous reviewers for their invaluable feedback on earlier versions of the article. The case study used in this article was funded by the Dutch Ministry of Social Affairs and Employment and the Ministry of Justice and Security.

Biographies

Guido Veldhuis is a Senior Scientist at the Defence, Safety and Security unit of TNO (The Hague, The Netherlands). He received his MSc from the Erasmus Mundus Program in System Dynamics. He develops methods and models to support decision-makers who face complex social, defence, and security-related challenges. Guido served as the chair of the BeNeLux chapter of the System Dynamics Society for several years.

Eefje Smits-Clijsen is a Senior Scientist and Consultant at the Defence, Safety and Security unit of TNO (The Hague, The Netherlands). She received her MSc in Criminology at the University of Leiden and a Master of Criminal Investigation at the Police Academy of the Netherlands. Before she started to work at TNO, she worked for the police and local governments. She learned to model in TNO projects, in which she now combines policy practice with modelling skills to support decision-makers with complex security-related challenges.

Rob van Waas is a researcher at TNO specialised in complex systems modelling applied to defence and security issues such as migration, climate security, and

DOI: 10.1002/sdr

defence-capability planning. In this role he enjoys the balance between research and application. He is a TU Delft Systems Engineering and Policy Analysis alumni and gained experience in different positions at L'Oréal in the Benelux before switching to a research role at TNO.

References

Andersen DF, Richardson GP. 1997. Scripts for group model building. *System Dynamics Review: The Journal of the System Dynamics Society* **13**(2): 107–129.
Black LJ. 2013. When visuals are boundary objects in system dynamics work. *System Dynamics Review* **29**(2): 70–86.
Bures V. 2017. A method for simplification of complex group CLDs based on endogenisation Encapsulation and Order-Oriented Reduction. *Systems* **5**(3): 1–22.
CBS 2021. Stapeling en (multi)problematiek tijdens de coronapandemie [Accumulation and (multiple) problems during the corona pandemic]. https://dashboards.cbs.nl/v3/stapeling_tijdens_de_coronapandemie/.
CBS 2022. *CBS Statline*. https://opendata.cbs.nl/#/CBS/nl/dataset/82242NED/table?ts=1661525483306.
Coyle G. 2000. Qualitative and quantitative modelling in system dynamics: Some research questions. *System Dynamics Review* **16**(3): 225–244.
Coyle G. 2001. Rejoinder to Homer and Oliva. *System Dynamics Review* **17**(4): 357–363.
de Gooyert V, Bleijenbergh I, Korzilius H, Fokkinga B, Lansu M, Raaijmakers S, Rouwette E, van der Wal M 2019. Why we do not always simulate. Retrieved 26 Aug. 2022 from: https://systemdynamics.org/why-we-do-not-always-simulate/.
de Klerk M, Plaisier I, Wagemans F. 2020. *Welbevinden Ten Tijde van Corona: Eerste Bevindingen op Basis van Een bevolkingsenquête uit Juli 2020 [Well-Being in the Time of Corona: First Findings Based on a Population Survey from July 2020]*. The Hague: Sociaal en Cultureel Planbureau. Retrieved 31 Aug. 2022 from: https://www.scp.nl/publicaties/publicaties/2020/09/10/welbevinden-ten-tijde-van-corona.-eerste-bevindingen-op-basis-van-een-bevolkingsenquete-uit-juli-2020.
Desthieux G, Joerin F, Lebreton M. 2010. Ulysse: A qualitative tool for eliciting mental models of complex systems. *System Dynamics Review* **26**: 163–192.
Directorate-general Society and COVID-19 [DGSC-19] 2021. Domeinoverstijgende herstelopgaven DG Samenleving en COVID-19 [Cross-domain recovery tasks DG Society and COVID-19]. Retrieved 31 Aug. 2022 from: https://www.rijksoverheid.nl/documenten/rapporten/2021/12/23/domeinoverstijgende-herstelopgaven-dg-samenleving-en-covid-19.
Doyle JK. 1997. The cognitive psychology of systems thinking. *System Dynamics Review: The Journal of the System Dynamics Society* **13**(3): 253–265.
Dutch Ministry of Finance 2021. Bijlage 5: Uitwerking Strategische Evaluatie Agenda [Appendix 5: Elaboration of the Strategic Evaluation Agenda]. Retrieved 31 Aug. 2022 from: https://www.rijksfinancien.nl/memorie-van-toelichting/2022/OWB/XV/onderdeel/1052062.
Eker S, Zimmermann N. 2016. Using textual data in system dynamics model conceptualization. *Systems* **4**(3): 1–14.
Engbersen G, van Bochove M, de Boom J, Etienne T, Krouwel A, van Lindert J, van Wensveen P 2021. De Ongeduldige Samenleving. De maatschappelijke impact van COVID-19 in Amsterdam, Den Haag, Rotterdam & Nederland. [The impatient society. The societal impact of COVID-19 in Amsterdam, The Hague, Rotterdam & the

DOI: 10.1002/sdr

Netherlands]. Retrieved from: https://www.narcis.nl/publication/RecordID/oai:hbokennisbank.nl:sharekit_hh%3Aoai%3Asurfsharekit.nl%3Ac6010874-d500-4b28-80b2-7aea7b33bf93.

Engzell P, Frey A, Verhagen MD. 2021. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* **118**(17): 1–7.

Farr WW, Allen SD, Tomoaia-Cotisel A, Hovmand PS. 2022. Documenting the modeling process with a standardized data structure described and implemented in DynamicVu. *System Dynamics Review* **38**(3): 264–291.

Forrester JW. 2007. System dynamics – The next fifty years. *System Dynamics Review* **23**: 359–370.

Gemeente Amsterdam 2020. Dashboard: Gevolgen corona voor Amsterdam [Dashboard: Consequences of corona for Amsterdam]. Retrieved 31 aug. 2022 from: https://onderzoek.amsterdam.nl/static/dashboard-corona/.

Ghaffarzadegan N, Lyneis J, Richardson GP. 2011. How small system dynamics models can help the public policy process. *System Dynamics Review* **27**(1): 22–44.

Ghaffarzadegan N, Rahmandad H. 2020. Simulation-based estimation of the early spread of COVID-19 in Iran: Actual versus confirmed cases. *System Dynamics Review* **36**: 101–129.

Homer J, Oliva R. 2001. Maps and models in system dynamics: A response to Coyle. *System Dynamics Review* **17**(4): 347–355.

Homer J. 2014. Levels of evidence in system dynamics modeling. *System Dynamics Review* **30**(1–2): 75–80.

Hovmand PS, Rouwette EAJA, Andersen DF, Richardson GP, Kraus A 2013. Scriptapedia 4.0.6. Retrieved from: http://www.systemdynamics.org/web.portal?P1405?0 on 22 Oct 2013.

Integraal Kankercentrum Nederland [IKNL] 2021. Aantal kankerpatiënten in 2020 gedaald door de coronacrisis, eerste daling in 30 jaar. [Number of cancer patients dropped in 2020 because of the coronacrisis, first drop in 30 years.]. Retrieved from: https://iknl.nl/persberichten/aantal-nieuwe-kankerpatienten-in-2020-gedaald-door#:~:text=in%20dertig%20jaar-,Aantal%20nieuwe%20kankerpati%C3%ABnten%20in%202020%20gedaald%20door%20coronacrisis%2C%20eerste%20daling,4.000%20ten%20opzichte%20van%202019.

Kim H, Andersen DF. 2012. Building confidence in causal maps generated from purposive text data: Mapping transcripts of the Federal Reserve. *System Dynamics Review* **28**: 311–328.

Kuipers EJ, van Gennip CEG 2022. Brief van de ministers van Volksgezondheid, Welzijn en Sport en van Sociale zaken en Werkgelegenheid [Letter from the ministers of Health, Welfare and Sport and of Social Affairs and Employment]. Retrieved 13 Aug. 2022 from: https://zoek.officielebekendmakingen.nl/kst-25295-1834.html.

Kuipers EJ, Adriaansens MAM, van Gennip CEG 2023. Brief van de ministers van Volksgezondheid, Welzijn, Economische zaken en Klimaat en Sport en van Sociale zaken en Werkgelegenheid [Letter from the ministers of Health, Welfare and Sport, Economic affairs and Climate policy and of Social Affairs and Employment]. Retrieved 13 Jan. 2023 from: https://www.tweedekamer.nl/downloads/document?id=2023d00620.

Lane DC. 2008. The emergence and use of diagramming in system dynamics: A critical account. *Systems Research and Behavioral Science* **25**(1): 3–23.

Lane DC. 1995. *The Folding Star: A Comparative Reframing and Extension of Validity Concepts in System Dynamics*. In proceedings of the international conference of the System Dynamics Society: Tokyo, Japan.

Luna-Reyes LF, Andersen DL. 2003. Collecting and analyzing qualitative data for system dynamics: Methods and models. *System Dynamics Review* **19**: 271–296.

DOI: 10.1002/sdr

Maani KE, Maharaj V. 2004. Links between systems thinking and complex decision making. *System Dynamics Review* **20**(1): 21–48.

Malbon E, Parkhurst J. 2022. System dynamics modelling and the use of evidence to inform policymaking. *Policy Studies* **44**(4): 454–472.

Martinez-Moyano IJ, Richardson GP. 2013. Best practices in system dynamics modeling. *System Dynamics Review* **29**(2): 102–123.

Morecroft JDW. 1982. A critical review of diagramming tools for conceptualizing feedback system models. *Dynamica* **8**(1): 20–29.

Niks IMW, Veldhuis GA, van Zwieten MHJ, Sluijs T, Wiezer NM, Wortelboer HM. 2022. Individual workplace well-being captured into a literature- and stakeholders-based causal loop diagram. *International Journal of Environmental Research and Public Health* **19**(15): 8925.

NSCR 2021. Stay Home or Stay Safe: results page. Retrieved 31 Aug. 2022 from: https://nscr.nl/stay-home-stay-safe-resultaten.

Plate R. 2010. Assessing individuals' understanding of nonlinear causal structures in complex systems. *System Dynamics Review* **26**(1): 19–33.

Richardson GP. 1999. Reflections for the future of system dynamics. *Journal of the Operational Research Society* **50**(4): 440–449.

Richardson GP. 1986. Problems with causal-loop diagrams. *System Dynamics Review* **2**(2): 158–170.

Richardson GP, Andersen DF. 1995. Teamwork in group model building. *System Dynamics Review* **11**(2): 113–137.

Richmond B. 1993. Systems thinking: Critical thinking skills for the 1990s and beyond. *System Dynamics Review* **9**(2): 113–133.

Richmond B. 1997. The "thinking" in systems thinking: How can we make it easier to master. *The Systems Thinker* **8**(2): 1–5 Retrieved 31 Aug. 2022 from: https://thesystemsthinker.com/the-thinking-in-systems-thinking-how-can-we-make-it-easier-to-master/.

RIVM 2021. Applying behavioural science to COVID-19. Overview of behavioural monitoring of COVID-19. Retrieved 31 Aug. 2022 from: https://www.rivm.nl/en/coronavirus-covid-19/research/behaviour.

Rouwette EAJA, Vennix JAM, van Mullekom T. 2002. Group model building effectiveness: A review of assessment studies. *System Dynamics Review: The Journal of the System Dynamics Society* **18**(1): 5–45.

Scott RJ, Cavana RY, Cameron D. 2016. Recent evidence on the effectiveness of group model building. *European Journal of Operational Research* **249**(3): 908–918.

Schaffernicht M. 2010. CLDs between structure and behaviour: A critical analysis of the relationship between polarity, behaviour and events. *Systems Research and Behavioral Science* **27**(6): 653–666.

Snel E, De Boom J, Van Bochove M, Engbersen G. 2021. Sociaal kapitaal als bescherming tegen de mentale gevolgen van COVID-19. [social capital as protection against mental consequences of COVID-19.]. *Mens & Maatschappij* **96**(2): 213–241.

Stave K, Hopper M. 2007. What constitutes systems thinking? A proposed taxonomy. In *Proceedings of 25th International Conference of the System Dynamics Society*: Boston, MA: System Dynamics Society.

Sterman JD. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill: Boston.

Struben J. 2020. The coronavirus disease (COVID-19) pandemic: Simulation-based assessment of outbreak responses and postpeak strategies. *System Dynamics Review* **36**(3): 247–293.

Trouw 2021. Zeker 68 patiënten in Tilburgs ziekenhuis kregen complicaties door uitstel van zorg [At least 68 patients in Tilburg hospital experienced complications due to

DOI: 10.1002/sdr

postponement of care]. [online] Trouw-Archive. Retrieved 31 Aug. 2022 from: https://www.trouw.nl/zorg/zeker-68-patienten-in-tilburgs-ziekenhuis-kregen-complicaties-door-uitstel-van-zorg~b7c5c457/.

Turner BL, Kim H, Andersen DF. 2013. Improving coding procedures for purposive text data: Researchable questions for qualitative system dynamics modeling. *System Dynamics Review* **29**: 253–263.

UWV 2020. Coronacrisis: impact op de werkgelegenheid verschilt per sector [Corona crisis: impact on employment differs per sector]. Retrieved 25 January 2021 from: https://www.uwv.nl/imagesdxa/coronacrisis-invloed-op-de-werkgelegenheid-verschilt-per-sector_tcm94-446745.pdf.

van Giessen A, De Wit A, Van Den Brink C, Degeling K, Deuning C, Eeuwijk J, Suijkerbuijk A 2020. Impact van de eerste COVID-19 golf op de reguliere zorg en gezondheid: Inventarisatie van de omvang van het probleem en eerste schatting van gezondheidseffecten. Retrieved 31 Aug. 2022 from: https://rivm.openrepository.com/handle/10029/624583.

Veldhuis GA, van Scheepstal P, Rouwette E, Logtens T. 2015. Collaborative problem structuring using MARVEL. *EURO Journal on Decision Processes* **3**(3–4): 249–273.

Veldhuis GA, Smits-Clijsen EM, van Waas RPM 2021. *De sociale impact van de coronacrisis [The social impact of the corona crisis]*. TNO 2021 R11223. The Hague: TNO. Available at: https://publications.tno.nl/publication/34638392/lXcphP/TNO-2021-R11223.pdf.

Vennix JA. 1996. *Group Model Building—Facilitating Team Learning Using System Dynamics*. Wiley: Chichester.

Vermeulen S, van Berkel S, Alink L. 2020. *Kindermishandeling tijdens de eerste lockdown*. Instituut Pedagogische Wetenschappen, Universiteit Leiden: Leiden Retrieved 31 Aug. 2022 from: https://www.huiselijkgeweld.nl/binaries/huiselijkgeweld/documenten/rapporten/2021/01/06/kindermishandeling-tijdens-eerste-lockdown/rapport-prevalentie-kindermishandeling-tijdens-covid-lockdown.pdf.

VNG. 2020. Werkgroep Sociale Impact van de Coronacrisis. *Eindverslag Werkgroep Sociale Impact van de Coronacrisis [Final report social impact of the Corona crisis by municipal working group]*. 1–73. Retrieved 31 Aug. 2022 from: https://vng.nl/sites/default/files/2020-05/eindverslag-werkgroep-sociale-impact-coronacrisis.pdf.

Wolstenholme EF. 1982. System dynamics in perspective. *Journal of the Operational Research Society* **33**(6): 547–556.

Wolstenholme EF. 1999. Qualitative vs quantitative modelling: The evolving balance. *Journal of the Operational Research Society* **50**(4): 422–428.

Zimmermann N, Pluchinotta I, Salvia G, Touchie M, Stopps H, Hamilton I, Kesik T, Dianati K, Chen T. 2021. Moving online: Reflections from conducting system dynamics workshops in virtual settings. *System Dynamics Review* **37**(1): 59.

DOI: 10.1002/sdr