PAR PUBLIC ADMINISTRATION REVIEW

Joris van der Voet
Leiden University, The Netherlands

Elvira Lems
BMC Advisory, The Netherlands

Decision-Makers' Generation of Policy Solutions amidst Negative Performance: Invention or Rigidity?

Research Article: HERBERT SIMON Symposium

Abstract: *Behavioral theory proposes that decision-makers engage in search to identify satisfactory policy solutions to organizational problems. For complex problems, required solutions cannot be identified merely through search, but must be created and designed. While behavioral theory suggests that negative performance will spur creative solution generation, threat-rigidity theory provides a competing view that decision-makers' creative efforts are ultimately thwarted through restricted information processing and constriction of control. We test these competing expectations through a survey-experimental comparison of the creativity of policy solutions in response to negative budgetary performance, generated by a nationally representative sample of local government decision-makers. The findings indicate that negative performance decreases the creativity of policy solutions as rated by policy experts, and reveals that key mechanisms of threat-rigidity theory are at play. This reduced generation of creative policy solutions amidst negative performance places limits on public organizations' capacity to adapt to adverse circumstances.*

Evidence for Practice

- This experiment examines decision-makers' capacity to generate creative policy solutions to complex societal problems, as rated by independent policy experts.
- More creative policy solutions are not seen as less practically or politically feasible and are rated as more effective, indicating that creativity is a crucial component of effective policy responses to wicked problems.
- Negative performance information decreases the creativity of policy solutions that decision-makers put forward.
- This reduced creativity is explained by threat-rigidity effects: decision-makers emphasize increased monitoring, reduced collaboration, and an efficiency orientation.

Behavioral theory (Simon 1997)—as initiated by *Administrative Behavior* (Simon 1997) and the Carnegie school more generally (Cyert and March 1992; March and Simon 1993)—has proven a pervasive and enduring influence on public administration research, theory and practice. Utilizing concepts such as performance relative to aspirations and problemistic search, behavioral theory asserts that organizational reform, innovation, and policy change are most in evidence in response to negative performance (Cyert and March 1992). In such processes of adaptation, Simon (1997, 77) places particular emphasis on the constrained cognitive capacity (or "bounded rationality") of organizational decision-makers. Decision-makers are not omniscient: In response to any problem, the full range of possible policy solutions and their consequences are not known to them. When faced with a particular problem, organizational decision-makers therefore engage in processes of search in order to identify a satisfactory solution (cf. Cyert and March 1992, 188).

For contemporary public administration research and practice, Simon's classic insights on administrative behavior are of continued relevance. However, the challenges that are faced by public administrations are increasingly complex, place multifaceted and contradictory demands on potential solutions, and have complex interdependencies with other societal problems (Moon 2020; Rittel and Webber 1973; Van der Wal 2020; Weber and Khademian 2008). Such "wicked problems," characterized by Simon (1997, 128) as ill-structured problems, pose a particular difficulty to organizational decision-makers. For complex societal problems, the required solution cannot be identified through search, but a novel solution must be created and designed. "The task is not to search but to synthesize: to design" (Simon 1997, 126). To address wicked societal problems, decision-makers' capacity to generate creative policy solutions is therefore of paramount importance.

In the commentary in the *Administrative Behavior*'s fourth edition, Simon (1997, 126) acknowledges that

Joris van der Voet is an Associate Professor of Public Management at the Institute of Public Administration, Leiden University. His research examines how decision-makers bring about change to enhance effectiveness, adaptivity, and resilience of public organizations in response to societal challenges, adverse performance, and declining financial resources. He is currently principal investigator on a research project that is funded by the Dutch National Science Foundation, titled "Administrative Attention amidst Political Failure."
Email: j.van.der.voet@fgga.leidenuniv.nl

Elvira Lems is a graduate (with honors) of the Master of Public Administration at Leiden University, with a specialization in Public Management and Leadership. Her master thesis addresses the effects of negative performance information on creativity of public sector decision-makers. Upon graduation, she was employed at Leiden University as an academic lecturer and researcher. She is currently a trainee at BMC Advisory, focusing on the policy fields of education and youth.
Email: elviralems@gmail.com

Public Administration Review, Vol. 82, Iss. 5, pp. 931–945. © 2022 The Authors. *Public Administration Review* published by Wiley Periodicals LLC on behalf of American Society for Public Administration.
DOI: 10.1111/puar.13462.

his concise treatment of the generation of alternatives is a "serious shortcoming" of the original text. This article explicitly focuses on the generation of alternatives phase of the behavioral model of decision-making, and specifically addresses the question to what extent negative performance evokes the generation of creative policy alternatives by organizational decision-makers. The article develops theoretical expectations based on behavioral theory and threat-rigidity theory. Based on behavioral theory, it is expected that negative performance will cause decision-makers to formulate more creative policy solutions, whereas threat-rigidity theory argues that decision-makers' creativity will be inhibited by negative performance. The research question is examined by means of a large-scale survey experiment in the context of youth care policy in Dutch municipalities, in which 996 municipal council members participated. Participants in the treatment condition receive negative budgetary performance feedback, and are asked to formulate up to five policy solutions to improve youth care implementation in their municipality. In total, 3,294 solutions were categorized into 31 categories through a content analysis. Subsequently, using Amabile's (1982) consensual assessment technique, aldermen with responsibility for youth care (n = 111) and directors of youth care organizations (n = 50) were invited to assess the creativity of the generated solutions.

This article makes two contributions to public administration research and theory. First, this article provides an empirical test of solution generation as a central yet empirically undocumented mechanism of behavioral theory. Prior empirical studies in public administration have documented how exposure to negative performance triggers strategic behavior as described in the theory. However, studies on the organizational level are unable to observe the intra-organizational mechanisms that are specified in the theory (Hong 2019; Rutherford and Meier 2015; Salge 2011; Zhu and Rutherford 2019). Studies on the individual level of analysis have provided evidence that decision-makers generally utilize performance information (Holm 2017, 2018) and update goal preferences in response to negative performance information (Nielsen 2014), but have not informed solution generation. This article provides experimental evidence on the effects of negative performance on solution generation based on multi-source data in a representative sample.

The second contribution of this article is that it tests the behavioral framework against expectations based on threat-rigidity theory. Prior empirical research on behavioral theory has resulted in inconsistent findings (Posen et al. 2017), which suggests that alternative mechanisms are at play (Ocasio 1997). We therefore incorporate threat-rigidity theory as a rival theoretical perspective on the generation of creative policy solutions in response to negative performance information. While *Administrative Behavior* acknowledges the relevance of stress, Simon (1997, 138) does not provide an account of how rational behavior of decision-makers may be further inhibited by threat. Threat-rigidity theory highlights the role of external stressors as constraints on decision-makers' strategic behavior, and states that organizational decision-makers' attempts to generate creative solutions will be ultimately thwarted by constraints on information processing capacities, constriction of control, and an efficiency-orientation brought on by mechanistic shift (Staw, Sandelands, and Dutton 1981).

The next section sets out the key theoretical principles of behavioral theory and threat-rigidity theory and derives the central hypotheses of the study. The empirical setting, as well as the procedures of the survey experiment, content analysis, and consensual assessment technique, are outlined in the section on methodology. In the Results section, we provide empirical evidence that negative budgetary performance reduces the generation of creative policy solutions. We complement our statistical analysis with a qualitative analysis that provides additional evidence that key mechanisms of threat-rigidity theory, most notably constriction of control and efficiency orientation, underlie this decreased creativity of solution generation in the face of negative performance. We then reflect on the continued relevance of *Administrative Behavior* for public administration research and practice, and discuss implications and limitations of this study.

Theory

Behavioral Responses to Performance below Aspirations

Behavioral theory portrays organizational decision-makers as rational actors that are intent on obtaining optimal returns on means-end relationships, but whose capacity to make fully rational decisions is ultimately limited (Simon 1997). Their behavior is informed by the performance information that provides cues about the degree to which current decisions are satisfactory, and can be called strategic in the sense that administrative decision-makers seek to remedy problems through the identification, selection, and implementation of more satisfactory policies (Cyert and March 1992). Behavioral theory proposes that decision-makers rely on a straightforward decision-rule to determine if a particular situation poses a problem, and as such requires decision-makers' attention and response. This decision-rule states that current levels of performance are compared to decision-makers' goals or *aspiration levels* (Cyert and March 1992; March and Simon 1993). Two types of aspiration levels are typically distinguished. First, performance falls short of *historical aspiration levels* when current the performance is lower than the organization's prior levels of performance. Second, performance is below *social aspiration levels* when the organization's performance is lower than the performance levels of relevant, comparable peer organizations (March and Simon 1993, 204).

Private sector research on performance below aspiration typically relies on quantitative measures of firm success such as return on assets or return on investment (e.g. Audia and Greve 2006). For public organizations, such indicators are often neither relevant nor available indicators of performance, as performance of public organizations is oftentimes difficult to conceptualize or measure (Andersen, Boesen, and Pedersen 2016). Nonetheless, prior public administration research has empirically examined behavioral responses to performance below historical and social aspiration levels, for instance by relying on measures of student achievement in the education sector (e.g. Holm 2017; Nielsen 2014; Nielsen and Moynihan 2017; Rutherford and Meier 2015), and hospital standardized mortality ratio or patient ratings in a health care setting (Salge 2011; Zhu and Rutherford 2019). These studies have examined behavioral responses ranging from innovative search (Salge 2011), prioritization (Holm 2017; Nielsen 2014), and managerial decision-making more generally (Hong 2019; Zhu and Rutherford 2019).

This study, which is set in a context of local government organizations, examines budgetary performance relative to historical and social aspiration levels. We define this as the degree to which the total spending of the municipality surpasses or falls within the

budget that was set for a given year. More specifically, negative historical budgetary performance refers to overspending the budget that was set in a given year, and negative social budgetary performance refers to overspending the budget more so than comparable peer organizations. For citizens as well as politicians on the local and national level, reform initiatives and policy discourse in the wake of the New Public Management have made public spending and cost-efficiency highly salient dimensions of public performance (Hvidman and Andersen 2015; van den Bekerom, Van der Voet, and Christensen 2020). Moreover, local government organizations face the formally set requirement to have a balanced annual budget, and are subject to oversight and face sanctions in case of severe or prolonged negative budgetary performance (Geißler, Hammerschmid, and Raffer 2019). In the context of local government organizations, budgetary performance is therefore likely a performance indicator to which organizational decision-makers are responsive, and research on financial cutback management and downsizing suggests negative budgetary performance may trigger creative responses characterized by increased risk-taking, learning, and collaboration (Kelman 2006; Van der Voet 2019).

When performance falls short of historical or social aspirations, behavioral theory proposes that organizational decision-makers engage in strategic behavior aimed at remedying the problem (Cyert and March 1992; March and Simon 1993). As decision-makers' rationality is limited by constraints on information processing, not all potential satisfactory policy solutions are known at any given point (Simon 1997). Behavioral theory thus states that decision-makers must therefore actively identify potential policy solutions by engaging in processes of search: the processing of information about urgent problems, alternatives, and their consequences (March and Simon 1993, 4). "Problemistic" search is driven by a problem—usually a specific one—and aimed at identifying a satisfactory policy solution to that problem (Cyert and March 1992, 169). In a public sector context, Salge (2011) provides empirical evidence that organizations engage in increased search in response to performance below aspirations. His study of public hospitals shows that research and development spending intensifies when patient mortality is above comparable peer organizations.

The Limits of Search: Solution Generation as a Creative Process

A recent literature review by Posen et al. (2017) indicates that theoretical expectations regarding search behavior are often not supported by empirical investigations. Simon (1997, 126) raises a more significant hesitation regarding the limits of problemistic search processes. Solution generation through problemistic search is incremental and small scale in nature (March 1991). However, decision-makers in the public sector operate in ill-defined problem landscapes, characterized by a complex, multilayered nature of societal problems, unclear means-ends relationships, and the unclear or politically contested requirements a potential solution must satisfy. For such wicked societal problems (Rittel and Webber 1973), bespoke policy solutions are therefore unlikely to be identified directly through processes of search. Instead of simply collecting alternative solutions, the task for decision-makers is not just to search but to synthesize distinct elements of potential solutions, thus generate rather than uncover potential policy solutions. This makes the generation of solutions essentially a *creative* process. Amabile (1982, 230) defines creativity as "the production of novel and appropriate ideas by individuals." In the literature on organizational innovation, creativity is seen as the process of generating novel ideas, and innovation refers to the subsequent stage of implementing ideas toward better procedures, practices, or products (Anderson, Potočnik, and Zhou 2014, 1298).

Simon does not portray creativity as an essentially distinct decision-making process other than the making of routine decisions (cf. Simon 1997, 128–32). Nonetheless, it is acknowledged that the generation of solutions to a complex problem "may involve inventing and elaborating whole performance programs where these are not already available in the problem solver's repertory" (March and Simon 1993, 161). As a key premise of behavioral theory, Cyert and March (Cyert and March 1992, 188) propose that "everything else being equal, relatively unsuccessful firms would be more likely to innovate than relatively successful firms." Recognizing that individual decision-makers' creativity underlies organizational innovation, we therefore propose that performance below historical or social aspirations increases decision-makers' generation of creative policy solutions. The following hypothesis is formulated:

> **Hypothesis 1:** Budgetary performance below historical or social aspirations positively affects decision-makers' generation of creative policy solutions.

An Alternative View: Threat-Rigidity Effects and Solution Generation

In setting out the psychology of administrative decision-making, Simon notes that stress may impact cognition to elicit counter-productive behavior, describing stress as "a powerful emotional force that can divert behavior from the urgings of reason" (Simon 1997, 138). In a situation of stress, the search processes that underlie solution generation may be "vigorous but largely stereotyped," leading to less fruitful solution generation (March and Simon 1993, 136). While *Administrative Behavior* and its intellectual descendants (Cyert and March 1992; March and Simon 1993) do not provide a comprehensive framework that accounts for the role of decision-making in stressful circumstances, such insights may be derived from threat-rigidity theory (Staw, Sandelands, and Dutton 1981). Threat-rigidity theory stipulates that decision-makers may perceive negative performance as a threatening situation, and that there may be a general tendency for individuals, groups, and organizations to behave rigidly in threatening situations (Staw, Sandelands, and Dutton 1981, 502). This rigidity is the result of two mechanisms: a restriction of information processing and a constriction of control.

Regarding the restriction of information processing, the theory echoes the abovementioned observations by Simon and March that threatening situations may distort search processes, thereby reducing the novelty and richness of the information that informs the generation of alternatives. "Even when search is increased, information received is likely to be similar to that of the past, due to heavy reliance on standard operating procedures, previous ways of understanding or communication that is low in complexity" (Staw, Sandelands, and Dutton 1981, 513). When under threat, decision-makers tend to focus on dominant cues and readily available information that supports their problem definition, such as information that is needed to justify or explain a response to a threat, while rejecting alternative information and cues

(Levine 1979). A case study by Deverell (2010) illustrates this mechanism, by showing how decision-makers react to threatening situations in line with existing routines and based on "prior expectations and internal hypotheses," rather than analyzing the current situation (Deverell 2010, 688).

In addition, threat-rigidity theory proposes that a constriction of control is likely to occur when decision-makers perceive negative performance as a threatening situation (Staw, Sandelands, and Dutton 1981, 513). In situations of threat, decision-making authority will become more centralized, and formalization and standardization will increase, bringing about a so-called mechanistic shift. The mechanistic organizational regime highlights an efficiency orientation—doing the same things better and cheaper—at the expense of flexibility and adaptation, and so contributes to a decreasing capability for creative policy responses. Several studies have found evidence of this mechanism in public organizations. Deverell (2010, 668) reports that amidst threat, "crucial decision regarding the prognosis was centralized to a few operational key functions." Wynen, Boon, Kleizen, and Verhoest (2020, 499) find that repeated structural reforms may trigger increased centralization and formalization on the managerial level, thereby "decreasing support for novel ideas generated by employees."

Threat-rigidity theory thus provides a rival view to behavioral theory, by stating that the generation of creative policy solutions may ultimately be thwarted by restricting and constraining the generation of creative policy responses. Based on this framework, we propose that performance below historical or social aspirations decreases decision-makers' generation of creative policy solutions. The following hypothesis is formulated:

> **Hypothesis 2:** Budgetary performance below historical or social aspirations negatively affects decision-makers' generation of creative policy solutions.

Methodology

Research Setting

The study's hypotheses are examined by means of a survey experiment in the context of Dutch youth care policy. Following a 2015 nationwide reform, youth care implementation in The Netherlands was decentralized from the national to the local government level. Dutch local government organizations—municipalities—are multipurpose organizations with a wide range of other policy responsibilities, which include social welfare, infrastructure planning and upkeep, environmental policies, cultural policy and sports, and the local economy. Dutch municipal organizations are highly homogenous in terms of institutional and structural characteristics, despite wide variation in population size (ranging from 947 to 872,229 inhabitants in 2020). The municipality council represents the municipality's legislative body, for which seats are assigned based on proportional representation. The college of mayor and aldermen is the municipality's executive body of the municipality. Aldermen have individual portfolios that typically include a range of policy domains.

For youth care policy, which is the policy domain examined in this study, municipalities operate on nationally set budgets and typically contract in services or personnel from privately-owned youth care organizations. Youth care services in the Netherlands consist of all forms of help and support to youths and their caregivers. Services vary from general and preventive care (for instance counseling and caregiver support) to intensive, highly specialized services such as forensic psychiatry. A major goal of the decentralization was to create more efficient and cost-effective youth care services, by making youth care provision more integral and reducing (access to) specialist care, as well as to stimulate innovative care provision by enabling more discretion and autonomy (Visser and Kruyen 2021).

Since the 2015 reform, youth care policy has developed into the premier policy issue on the local government level. Youth care policy has proven to be "wicked" policy problem, as it is concerned with unclear means-ends relationship (Simon 1997). Potential policy solutions are therefore subject to high levels of uncertainty, as well as conflicting preferences along progressive and conservative political orientations. Moreover, complex interdependencies with other policy fields exist, including education, housing, and safety. A widespread presence of overlapping and conflicting national laws, municipality policy, and procedures of youth care organizations amount to severe administrative burdens and inefficiencies, which the 2015 reform intended to address but failed to resolve (ZonMW 2018). Most importantly, youth care policy takes place in a context of severe budget shortages (Binnenlands Bestuur 2020), as the main objective of the policy decentralization was to increase efficiency and cost-effectiveness of youth care implementation. The 2015 reform resulted in a yearly budget reduction of 15% (the total budget is estimated at 3.459 billion euro in 2018), while demand for youth care services has seen yearly increases. As a result, 92.3% of Dutch municipalities overspent their youth care budget in 2018. For this reason, the provision of budgetary performance information is highly salient for decision-makers whose responsibility is approving and overseeing the municipality's budget. A second consideration is that youth care performance is inherently subjective and difficult to quantify (Nooteboom et al. 2019; Rap et al. 2019). In addition, the decentralization reform has increased the fragmentation of performance measurement, further limiting the extent to which performance can systematically be compared across municipalities. Whereas prior public administration research in the behavioral tradition has examined different indicators of performance (e.g. Nielsen 2014; Salge 2011; Zhu and Rutherford 2019), we believe this study's emphasis on budgetary performance information is theoretically warranted and practically salient.

Although the organizational structure and institutional design of the political-administrative component of Dutch municipalities is highly homogenous, municipalities differ strongly in terms of their socio-economic and demographic characteristics, as well as in terms of political preferences. This limits the degree to which best-practice policy solutions can effectively travel across municipal borders. Youth care policy is thus a prime example of Simon's (1997, 128) ill-structured problem, and places demands on decision-makers' capacity to design bespoke policy solutions.

Sample

A survey experiment was conducted among local government council members in a five-week period, starting November 212,019. Local council members are an appropriate decision-maker in this context, as the municipality council collectively approves the

budget for youth care spending and decides on the way youth care policy is implemented. The total Dutch population of municipality councilors consists of 8,261 individuals, making them the largest group of elected political decision-makers in The Netherlands. An invitation to participate in the survey was sent to the email addresses, collected from municipality websites, of 7,815 local council members in 328 out of 355 Dutch municipalities. Data collection resulted in a complete response of 996 participants who have fully participated in the survey-experiment (response rate: 12.7%), representing 301 out of 355 municipalities (84.8% of the municipalities in The Netherlands). This response rate is in line with recent studies in comparable contexts (i.e. George et al. 2020; Meyfroodt and Desmidt 2020; Van der Voet 2021).

Concerning the representativeness of the sample, we note that 27 municipalities with fewer than 30,000 inhabitants could not be included in the study because individual contact information of council members and/or information regarding youth care spending was unavailable. Comparing the characteristics of the sample to the total population of municipality council members (VNG 2020), we conclude that the sample is representative in terms of female gender proportion ($\bar{X}$ = .32, μ = .31, p = .33). The sample is slightly younger than the population ($\bar{X}$ = 54.7, μ = 52.7, p < .00). In terms of main political party representation, the sample does not differ from the population in terms of membership of *VVD*,[1] *D66*[2] and *PvdA*[3] political parties. The Christian-conservative *CDA* party and local parties are slightly underrepresented in the sample, while the green *Groenlinks* party is slightly overrepresented. The sample thus appears representative in terms of age, partly representative in terms of party membership, and slightly more representative of younger than older council members.

Research Procedures and Measures

Figure 1 details the procedures of the study, which consists of a survey experiment, a content analysis, and the application of the consensual assessment technique.

Survey Experiment. The study relies on a survey experimental design to communicate performance information to participants of the study and to collect their policy solutions. The flowchart in Appendix C depicts the procedures of the survey experiment. First, we distinguish between participants based on natural variation of budgetary performance. As the study is concerned with overspending, 73 respondents (from 21 different municipalities) are excluded from the experiment as their municipality did not overspend the youth care budget. Further, we differentiate between municipalities that have lower levels of overspending than the average of municipalities that are comparable in size (i.e. higher budgetary performance; 156 municipalities) and municipalities that have higher levels of overspending than comparable peers[4] (i.e. lower budgetary performance; 124 municipalities).

Second, participants within both categories are randomly assigned to either control or treatment condition. The experiment uses experimental vignettes in line with prior experiments on performance information (Desmidt and Meyfroodt 2021; Nielsen and Moynihan 2017). The baseline vignette introduces respondents to the topic of youth care budgeting. For respondents in the control conditions (group 1 and group 3), this information is complemented with general municipality-unspecific information. For the treatment condition, the baseline vignette is complemented with performance information that signals budgetary performance below *historical* aspirations for respondents in municipalities with lower levels of overspending (group 2). Respondents in municipalities with higher levels of overspending receive additional performance information that signals budgetary performance below *social* aspirations when assigned to the treatment condition (group 4). The full experimental vignettes are given in Appendix C.

Third, participants are asked to formulate up to five policy solutions to improve youth care implementation in their municipality: *"In your view, what are opportunities to further change and improve youth care in your municipality? We ask you to write down as many ideas as possible."* This resulted in the generation of 3,294 unique policy solutions.

Content Analysis. The generated solutions were categorized by means of a content analysis. All solutions were read to make an initial list of the most frequent topics addressed in the solutions. Subsequently, usable solutions were combined together into solution categories based on their topic in an iterative process. Utilizing an iterative process and a pilot test, 31 categories of solutions were formed that together describe the topics that were expressed by decision-makers. All solution categories are given in Appendix A, and Appendix B provides an overview of the categorization process of the content analysis. Responses that were too vague, unintelligible, formulated as rhetorical question or problem-definition, or that could otherwise not be categorized were excluded. In total, 2,812 usable solutions were categorized. The most frequently mentioned solutions by decision-makers were "reducing the intake of clients in youth care" (349 times), "a more integral approach to youth care policy" (232 times), "more emphasis on low-intensity, generalist care" (208 times), and "more attention for early detection and prevention of problems" (203 times).

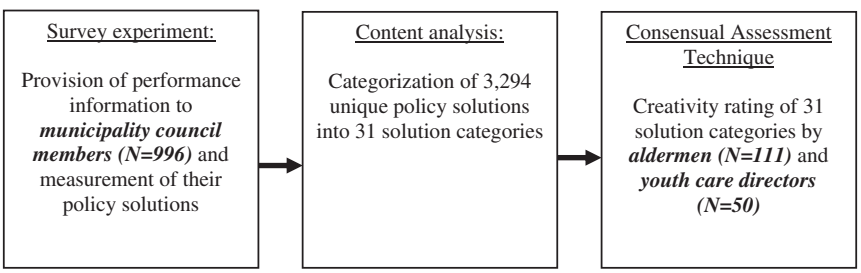

Figure 1 Overview of the Study's Procedures

Consensual Assessment Technique. The creativity of all 31 solution categories was measured using Amabile's (1982) consensual assessment technique. This technique departs from the observation that the measurement of creativity is characterized by a criterion problem: Although conceptual definitions of creativity exist in the academic literature, there is a "lack of a clear operational definition and an appropriate assessment methodology" (Amabile 1982, 997). The consensual assessment technique relies on the view that a solution may be judged as "*creative to the extent that appropriate observers independently agree it is creative. Appropriate observers are those familiar with the domain in which the product was created or the response articulated. Thus, creativity can be regarded as the quality of products or responses judged to be creative by appropriate observers* […]." (Amabile 1982, 1001). Based on the consensual assessment technique, it is assumed that "it is possible to obtain reliable judgments of creativity" of policy solutions and that experts agree on this perception of creativity with each another (Amabile 1982, 1001). A solution is thus creative when independent experts perceive it to be creative.

As youth care policy on the local level depends on both policy formulation as well as policy implementation, we rely on two sets of experts to provide to rate the creativity of policy solutions from two different perspectives. First, a sample of municipality aldermen was recruited to provide a political perspective on youth care policy. Whereas mayors have representative function as head of the municipality but have few policy responsibilities other than public safety, aldermen form the executive power of the municipality. Each municipality typically has a single alderman with responsibility for the youth care portfolio. Municipality aldermen did not partake in the original experiment in which policy solutions were generated, as aldermen are not simultaneously members of the municipal council in the Dutch context. An invitation to participate in an expert survey was distributed to 355 aldermen across all Dutch municipalities on May 25, 2020. Second, a sample of directors of youth care organizations was recruited, as they have a professional perspective with particular expertise about youth care implementation. A total number of 128 youth care organizations were contacted. After 1 week, a reminder was sent. After 2 weeks this resulted in 50 responses of individual youth care directors and 111 responses of aldermen.

The expert survey presented all 31 solution categories as described in Appendix A in random order, and respondents were asked to rate the solutions on a scale from 1 to 10 on three indicators: creativity (introduced to respondents as "*The degree to which the idea is novel and original*"), effectiveness ("*The degree to which you expect the idea to be efficacious*"), and feasibility ("*The degree to which the idea is practically and politically achievable*"). The inclusion of indicators other than creativity is a recommended practice by Amabile (1982), as it allows to examine to what extent assessments of creativity are empirically distinct from other criteria.

The expert ratings for all solution categories were assigned to all generated policy solutions, which were then aggregated the level of individual decision-makers as average creativity, effectiveness, and feasibility ratings. A key assumption of the consensual assessment technique is that experts should be able to agree with another on the degree of creativity. An examination of internal consistency of the experts' creativity ratings shows that this assumption is met, with satisfactory values for Cronbach's alpha (α = .927 aldermen's creativity assessments; α = .942 for youth care directors' creativity assessment). These values indicate that the creativity assessments of both expert panels are highly reliable. The average creativity ratings for the two samples of independent judges are reported in Appendix A.

Balance Check

We perform a balance check to examine random assignment of participants to treatment and control conditions by performing t-tests for participants gender, age (in years), tenure (in years), political orientation (measured on scale from "1: Strong left" to "5: Strong right"), membership of coalition vs. opposition, number of solutions generated, and the size of the youth care budget shortage in the participant's municipality (in percentages of the total budget). The results are given in table 1. The results indicate that no statistically significant differences exist between the treatment groups vis-à-vis their relative control group concerning age, tenure, political orientation, number of ideas generated, and the size of the budget shortage. For gender, the proportion of females in group 3 is lower than group 4 (difference of .10, p = .04).

Results

The expert panels of aldermen and youth care directors scored all solution categories on the dimensions creativity, effectiveness, and feasibility. The mean values and standard deviations of the experts' ratings are given in table 2. The results indicate that on average, the experts perceive the respondents' solutions to be only moderately creative. Aldermen rate the average creativity of the formulated solutions with 5.8 on a ten-point scale. Directors of youth care organizations evaluate the creativity of the solutions with an average score of 5.4. The evaluations of effectiveness and feasibility are notably higher than the evaluation of creativity. The standard deviations indicate that ratings of effectiveness

Table 1 Balance Check and Descriptive Statistics

	Budgetary Performance Below Historical Aspirations			Budgetary Performance Below Social Aspirations			
Variable	Group 1: Control (N = 243)	Group 2: Treatment (N = 273)	p Value	Group 3: Control (N = 197)	Group 4: Treatment (N = 210)	p Value	Full Sample
Female gender	.36 (.48)	.31 (.46)	.19	.26 (.44)	.36 (.48)	.04	.32 (.47)
Age	53.7 (12.6)	54.6 (13.0)	.43	56.0 (12.2)	54.5 (12.3)	.25	57.4 (12.6)
Tenure	5.9 (5.8)	6.5 (6.8)	.26	6.7 (6.3)	6.3 (6.5)	.60	6.5 (6.5)
Political orientation	2.8 (.9)	2.8 (.8)	.67	2.7 (.9)	2.9 (.9)	.29	2.8 (.9)
Coalition membership	.56 (.5)	.54 (.5)	.76	.59 (.5)	.62 (.5)	.49	.57 (.4)
Number of solutions generated	2.9 (1.5)	3.1 (1.4)	.36	3.0 (1.4)	3.2 (1.4)	.15	3.0 (1.4)
Budget shortage	17.8 (10.7)	18.3 (10.4)	.62	61.1 (40.1)	60.2 (51.4)	.85	33.7 (38.4)

and feasibility are subject to higher variation than the ratings of creativity. This underlines that there is a high level of consistency in experts' creativity ratings. The dispersion of the three dimensions is further described by means of boxplots in figure 2. The boxplots show that for creativity, all alderman ratings of creativity range between 4.4 and 6.7 on a ten-point scale. The youth director ratings of creativity range between 3.9 and 6.8. Correlation analysis shows that creativity and effectiveness are strongly correlated (r = .728; p = .000; N = 62 solution categories) and that there is a weak, positive correlation between creativity and feasibility (r = .210; p = .102; N = 62 solution categories). This correlation is however not statistically significant.

Regression analysis is used to test the study's hypotheses. Individual decision-makers are nested within municipalities with unobserved characteristics that may confound the relationship of interest. We therefore examine the relationship between negative budgetary performance and creativity by means of hierarchical linear regression analysis with municipality fixed effects. In such a way, the statistical analysis controls for unobserved variation in characteristics on the municipality level. As random assignment to control and treatment conditions was successful, we do not model additional control variables on the level of individual respondents. We model dummy variables for the two experimental treatments. The dummy variable "budgetary performance information below historical aspirations" is used to model a comparison between group 1 and group 2. The dummy variable "budgetary performance information below social aspirations" is used to model a comparison between group 3 and group 4. In table 3 (performance below historical aspirations; group 1 and 2) and table 4 (performance below social aspirations; group 3 and 4), we report the unstandardized regression coefficient (b) and standard error in parentheses, the standardized regression coefficient (β), t-value (t), and p-value for statistical significance (p). The F-value is reported as an indicator of overall model fit.

Table 2 Central Tendency and Dispersion of Expert Ratings

Dimension	Aldermen		Directors	
	Mean	SD	Mean	SD
Creativity	5.803	0.784	5.435	0.786
Effectiveness	7.042	1.172	6.841	1.178
Feasibility	6.595	1.109	6.405	0.991

The results of the regression analysis in table 3 indicate that performance below historical aspirations negatively affects the creativity ratings of aldermen. The creativity of generated solutions in the experimental group (group 2) is thus lower than in the relative control group (group 1). The effect size of the identified effects of the experimental treatment accounts for roughly .2 standard deviations of aldermen's creativity ratings, which indicates a small—but not trivial—effect size. This result provides support for hypothesis 2, and in contradiction of hypothesis 1. For creativity ratings by youth care directors, however, no negative effect is present. Although the direction of the relationship is negative, the effect is not statistically significant. The results given in table 4 mirror those of table 3. Budgetary performance below social aspirations negatively affects the creativity ratings by aldermen, but not the creativity ratings of youth care directors.[5] Overall, we conclude that hypothesis 1 is rejected, and that the analysis provides partial support for the hypothesis that budgetary performance below historical and social aspirations reduces the generation of creative solutions (hypothesis 2).

Analysis of Solution Content: Theoretical Mechanisms

The regression analysis provides statistical evidence that negative budgetary performance decreases the creativity of generated

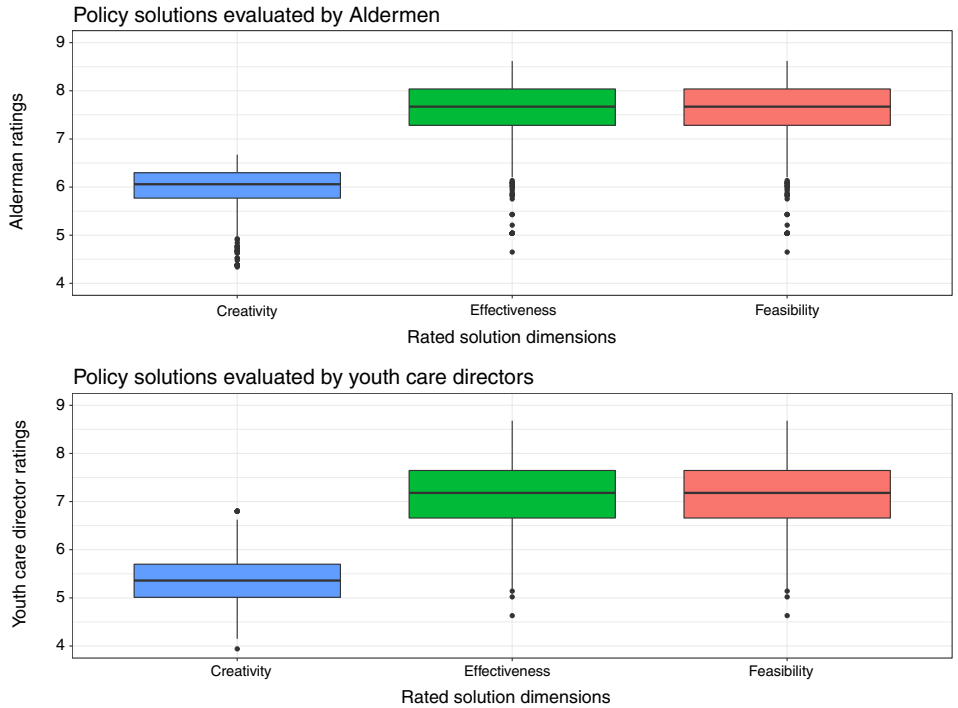

Figure 2 Boxplot of Policy Solution Evaluations by Alderman and Directors

Table 3 Regression Analysis for Budgetary Performance below Historical Aspirations (Group 1 and 2)

	Aldermen Creativity Ratings				Youth Care Director Creativity Ratings			
	b	*β*	*t*	*p*	*b*	*β*	*t*	*p*
Intercept	6.02 (.03)		189.63	.00	5.41 (.03)		156.83	.00
Dummy variable: Budgetary performance below historical aspirations	−.09* (.04)	−.19	−2.08	.04	−.05 (.04)	−.10	−1.14	.26
F value		4.30				1.29		
N				155 municipalities				
n				498 decision-makers				

*Statistical significance with $p < .05$.

Table 4 Regression Analysis for Budgetary Performance below Social Aspirations (Group 3 and 4)

	Aldermen Creativity Ratings				Youth Care Director Creativity Ratings			
	b	*β*	*t*	*p*	*b*	*β*	*t*	*p*
Intercept	6.04 (.03)		196.84	.00	5.36 (.04)		150.11	.00
Dummy variable: Budgetary performance below social aspirations	−.09* (.04)	−.22	−2.19	.03	−.00 (.04)	−.00	−.03	.98
F value		4.79				.00		
N				124 municipalities				
n				398 decision-makers				

*Statistical significance with $p < .05$.

policy solutions as perceived by aldermen. The direction of this relationship is thus in support of threat-rigidity theory (Staw, Sandelands, and Dutton 1981). Threat-rigidity theory puts forward observable theoretical mechanisms, including a restriction of information processing, a constriction of control and an efficiency orientation. In this section, we provide a qualitative analysis of differences in solution content between the treatment groups and their relative control groups, in order to examine to what extent the theoretical mechanisms of threat-rigidity are at play.

Utilizing t-tests to compare the frequency in which all policy solution categories are put forward, we first find evidence that negative budgetary performance brings about a constriction of control. In response to budgetary performance information below social aspirations (group 4), respondents are significantly more likely to propose "increased monitoring of the quality of youth care organizations" as a solution (solution category 23; $p = .018$). For example, decision-makers put forward solutions such as "To investigate how time is being spent," "To check operations on the work floor more often," and "To formulate, measure and report key performance indicators." Such measures are prime examples of enhancing control in the organization, which theoretically can be understood as the importance of organizational action increases under threat (Staw, Sandelands, and Dutton 1981, 513). In addition, participants that have received budgetary performance information below historical aspirations (group 2) are less likely to formulate solutions that emphasize the participation of youths, parents, and citizens in general (solution category 27, $p = .052$). Examples of such solutions are "To increase consultation with youths and parents," and "to initiate conversations with youths in an accessible way." The absence of consultation-based solutions is indicative of centralization, in which the number of decision-making participants is restricted (Staw, Sandelands, and Dutton 1981, 514).

Second, we find evidence that negative budgetary performance brings about an efficiency orientation, indicated by a greater dominance of concerns related to budgets, cost-cutting, and fine-tuning current operations. Participants in both the historical aspirations treatment (group 2; $p = .083$) and the social aspirations treatment (group 4; $p = .045$) are more likely than their counterparts to propose cutting back on other policy domains so that more funds remain available for youth care (solution category 29). For instance, decision-makers propose "To screen other municipal budgets to find budgetary room for youth care," and "To not cut back on youth care but on other concerns." Moreover, decision-makers in the historical aspirations treatment (group 2) more often propose to request additional financial resources for youth care as a municipality (solution category 10; $p = .020$), and decision-makers in the social aspirations group (group 4) are less likely to emphasize more research and innovation as a policy solution (solution category 26, $p = .018$). Both these solutions are indicative of threat-rigidity's mechanism of efficiency orientation, and the decreased reliance on research and innovation appears to be in direct opposition to behavioral theory's mechanism of failure-induced search.

This analysis thus reveals that the mechanisms of threat-rigidity theory underlie decreased creativity in response to negative budgetary feedback: decision-makers' policy solutions become more reflective of constriction of control and efficiency orientation, which are in turn rated as less creative solutions by aldermen and youth care directors.

Discussion

Seventy-five years after the publication of *Administrative Behavior*, behavioral theory and the Carnegie school on decision-making remain pervasive in public administration research. The theory's emphasis on strategic behavior in response to performance shortfalls and adversity more generally is often invoked in empirical research

to argue that public organizations initiate strategic changes and innovation in response to negative performance (Holm 2018; Hong 2019; Nielsen 2014; Zhu and Rutherford 2019). As decision-makers' knowledge of possible policy solutions is incomplete, this study has examined behavioral theory's assertion that the generation of creative policy solutions will be more in evidence in response to negative performance. We have contrasted behavioral theory with threat-rigidity theory, which instead states that creativity will be less in evidence as a result of restricted cognitive processes.

Based on a survey experimental design and the analysis of more than 3,000 policy solutions, we conclude that the generation of creative policy solutions is not increased in response to negative performance. We also do not find evidence of increased effort of decision-makers to generate policy solutions (Salge 2011), as indicated by the number of ideas generated (see table 1), their time spent generating solutions, or the length of generated solutions.[6] As such, we conclude that this study's findings are not in line with behavioral theory. Instead, we find partial statistical support for threat-rigidity theory's assertion that solution generation is inhibited amidst negative performance. Our qualitative analysis provides additional support for threat-rigidity theory by showing that generated policy solutions in the treatment groups are increasingly subject to constriction of control and efficiency orientation.

The study's results are subject to two important caveats. First, the effect size of the identified relationship is relatively modest, which is likely due to the survey experimental design of the study. Although the experiment has relied on factual performance information and real-world decision-makers, the magnitude of the effect may be understated due to the controlled and isolated setting in which the experimental task took place. The survey experiment can be considered a conservative test to identity threat-rigidity effects because participants are not exposed to elements of a real decision-making setting that bring on perceived threat, such as accountability, re-election concerns, and real-life consequences for municipal inhabitants. We acknowledge that although qualitative data were collected in the form of policy solutions, qualitative methods such as interviews may be more suitable than open-ended survey questions to empirically observe threat-rigidity effects. We therefore believe it is likely that the identified threat-rigidity effect generalizes to real-world decision-making settings, and that the effect size may be larger in magnitude.

A second caveat is that the results provide only partial support for hypothesis 2, as a decrease in creativity is observed based on the creativity ratings of municipality aldermen (politicians) but not youth care directors (managers). For youth care directors, the direction of the effect is also negative but it is not statistically significant. This divergence generally underlines the observation that creativity is in the eye of the beholder (Amabile 1982). Based on this study, we cannot fully explain why the creativity perceptions of politicians differ from managers in the way they do. In particular for ill-structured problems such as youth care, the assessment of what constitutes creativity depends on the experience, judgment and organizational role of individual decision-makers (cf. Simon 1965, 33). In assessing creativity, it is possible that politicians take a more integral perspective of a policy solution's coherence with overall municipality policy and base their creativity ratings on ideological assessment of the political suitability and legitimacy of policy solutions, while managers base their assessment more strongly on substantive knowledge, technical expertise and expected efficacy of policy solutions. For instance, based on their classic comparison of political and managerial decision-makers, Aberbach et al. (1990, 3) state that "[b]ureaucrats tend to emphasize the technical and de-emphasize the advocacy and partisan considerations within their role understanding, whereas politicians tend to do exactly the reverse." Comparative investigation of the creativity assessments of political and managerial actors thus constitutes a future research avenue for public administration research.

Our qualitative analysis identifies several mechanisms of threat-rigidity theory in idea generation behavior, and thus strengthens confidence that decision-makers' efforts to generate creative solutions is ultimately outweighed by routinized behaviors brought on by threat (cf. Kelman 2006). When decision-makers perceive negative performance information as threatening (Jordan and Audia 2012), their rational behavior is inhibited through mechanisms that limit information complexity and elaboration processes that underlie creativity (Deverell 2010; Staw, Sandelands, and Dutton 1981). The results suggest that – especially when negative performance is perceived as threatening – constructive responses may give way to self-enhancement (Jordan and Audia 2012), blame avoidance (Nielsen and Moynihan 2017) or negative emotional responses such as anger (Hattke, Hensel, and Kalucza 2020). Threat can serve as a boundary condition in such a way that behavioral theory applies more straightforwardly when perceived threat of performance information is low (cf. Jordan and Audia 2012). As alluded to by Simon (1997, 138), complementing behavioral theory with perspectives that inform when negative performance is perceived as threatening thus holds much promise for a more comprehensive insight in administrative behavior.

A limitation of the research design is that it does not allow to compare the effect of performance below historical aspirations to the effect of performance of social aspirations, as other public administration research has done (e.g. Nielsen 2014; Zhu and Rutherford 2019). A limitation of experiments that communicate performance information to decision-makers is that the vignettes provided to control and treatment conditions are dissimilar in length (cf. Desmidt and Meyfroodt 2021; Nielsen and Moynihan 2017). Although this study addresses this issue by complementing the vignette of the control condition with municipality-unspecific information. A potential threat to internal validity of the results remains, as the complementary information in the control vignette does not contain performance information. In examining the creativity of policy solutions, we have studied decision-makers in a particular policy field and national setting. As a result, the generalizability of the findings beyond this setting may not be warranted. Future research can utilize the combination of an experimental design in which salient performance information is provided with the consensual assessment technique based on expert ratings to examine the generation of creative policy solutions in different contexts.

Our investigation of creativity amidst negative budgetary performance has several implications for public administration practice. First, while several national public administrations have

used proportional budget reductions to incentivize increased innovation on lower levels of government, this study provides evidence that local government decision makers' creativity is reduced in amidst financial resource scarcity. As individual creativity is a critical determinant of organizational innovation, policy makers should critically consider the efficacy of incentivizing innovation through fiscal squeeze. A second implication for practice is that decision-makers should be aware that threat-rigidity effects can stifle creativity in decision-making processes, particularly when negative performance information is perceived to be threatening. Awareness that such effects occur through an increased emphasis on monitoring and control, as well as attempts to enhance efficiency and save costs, can help decision-makers to recognize threat-rigidity effects in decision-making processes. When creativity in solution generation is required, decision-makers can attempt to counter threat-rigidity effects by utilizing more divergent information and diverging from routinized decision-making procedures. A third implication for practice is that creativity is subjective and in the eye of the beholder. The policy solutions of council members are only moderately perceived to be creative by the expert panels, and political and managerial experts rate the creativity of individual policy solutions differently. Decision-makers in need of creative solutions would thus do well to probe the perceived creativity of potential policy solutions among myriad of actors, including political overseers, public managers, front-line implementors, partner organizations, and clients (cf. Visser and Kruyen 2021). In such a way, the consensual assessment technique can be applied in practice to measure and strengthen the creativity of policy solutions used to address complex societal problems.

Conclusion

The complex challenges that are faced by contemporary public administrations increasingly require novel and creative policy solutions. While behavioral theory suggests that decision-makers may generate satisfactory policy solutions to negative performance through processes of search, this article shows that the generation of creative policy solutions is reduced by negative performance. In the policy solutions generated by actual decision-makers in an experimental setting, we find evidence of threat-rigidity effects that underlie a reduced generation of creative policy solutions. For continued relevance of *Administrative Behavior*, future public administration research may benefit from increased attention for routinized cognitive and emotional responses that may inhibit rational responses to negative performance.

Notes

1. Volkspartij voor Vrijheid en Democratie: Party for Freedom and Democracy (conservative-liberal political party).
2. Democraten 66: Democrats 66 (social liberal political party).
3. Partij van de Arbeid: Labour Party (social-democratic political party).
4. Peer organizations were identified based on the number of inhabitants, using six classes: (1) fewer than 20.000 inhabitants, (2) 20.000 to 30.000 inhabitants, (3) 30.000 to 50.000 inhabitants, (4) 50.000 to 100.000 inhabitants, (5) 100.000 to 250.000 inhabitants, and (6) more than 250.000 inhabitants.
5. For both tables 3 and 4, *Z*-tests indicate that the differences in regression coefficients between aldermen and youth care directors are not statistically significant.
6. Additional analyses available with first author on request.

References

Aberbach, Joel D., Hans-Ulrich Derlien, Renate Mayntz, and Bert A. Rockman. 1990. American and German Federal Executives-Technocratic and Political Attitudes. *International Social Science Journal* 42(123): 3–16.

Amabile, Teresa M. 1982. Social Psychology of Creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology* 43(5): 997–1013. https://doi.org/10.1037/0022-3514.43.5.997.

Andersen, Lotte Bøgh, Andreas Boesen, and Lene Holm Pedersen. 2016. Performance in Public Organizations: Clarifying the Conceptual Space. *Public Administration Review* 76(6): 852–62. https://doi.org/10.1111/puar.12578.

Anderson, Neil, Kristina Potočnik, and Jing Zhou. 2014. Innovation and Creativity in Organizations: A State-of-the-Science Review, Prospective Commentary, and Guiding Framework. *Journal of Management* 40(5): 1297–333. https://doi.org/10.1177/0149206314527128.

Audia, Pino G., and Henrich R. Greve. 2006. Less Likely to Fail: Low Performance, Firm Size, and Factory Expansion in the Shipbuilding Industry. *Management Science* 52(1): 83–94. https://doi.org/10.1287/mnsc.1050.0446.

van den Bekerom, Petra, Joris van der Voet, and Johan Christensen. 2020. Are Citizens more Negative about Failing Service Delivery by Public than Private Organizations? Evidence from a Large-Scale Survey Experiment. *Journal of Public Administration Research and Theory* 31: 128–49. https://doi.org/10.1093/jopart/muaa027.

Binnenlands Bestuur. 2020. "Onderzoek wijst uit: kosten jeugdzorg onhoudbaar." Binnenlands Bestuur. 2020. https://www.binnenlandsbestuur.nl/sociaal/nieuws/gemeenten-komen-1-6-tot-1-8-miljard-tekort-op.15476038.lynkx [accessed December 18, 2020].

Cyert, Richard M., and James G. March. 1992. *Behavioral Theory of the Firm*. Cambridge, MA.

Desmidt, Sebastian, and Kenn Meyfroodt. 2021. How Does Public Disclosure of Performance Information Affect Politicians' Attitudes towards Effort Allocation? Evidence from a Survey Experiment. *Journal of Public Administration Research and Theory* 31(4): 756–72.

Deverell, Edward. 2010. Flexibility and Rigidity in Crisis Management and Learning at Swedish Public Organizations. *Public Management Review* 12(5): 679–700. https://doi.org/10.1080/14719031003633946.

Geißler, René, Gerhard Hammerschmid, and Christian Raffer. 2019. "Local Public Finance in Europe." Bertelsmann Stiftung, 252.

George, Bert, Martin Baekgaard, Adelien Decramer, Mieke Audenaert, and Stijn Goeminne. 2020. Institutional Isomorphism, Negativity Bias and Performance Information Use by Politicians: A Survey Experiment. *Public Administration* 98(1): 14–28.

Hattke, Fabian, David Hensel, and Janne Kalucza. 2020. Emotional Responses to Bureaucratic Red Tape. *Public Administration Review* 80(1): 53–63. https://doi.org/10.1111/puar.13116.

Holm, Jakob Majlund. 2017. Double Standards? How Historical and Political Aspiration Levels Guide Managerial Performance Information Use. *Public Administration* 95(4): 1026–40. https://doi.org/10.1111/padm.12379.

———. 2018. Successful Problem Solvers? Managerial Performance Information Use to Improve Low Organizational Performance. *Journal of Public Administration Research and Theory* 28(3): 303–20. https://doi.org/10.1093/jopart/muy017.

Hong, Sounman. 2019. A Behavioral Model of Public Organizations: Bounded Rationality, Performance Feedback, and Negativity Bias. *Journal of Public Administration Research and Theory* 29(1): 1–17. https://doi.org/10.1093/jopart/muy048.

Hvidman, Ulrik, and Simon Andersen. 2015. Perceptions of Public and Private Performance: Evidence from a Survey Experiment. *Public Administration Review* 76(September): 111–20. https://doi.org/10.1111/puar.12441.

Jordan, Alexander H., and Pino G. Audia. 2012. Self-Enhancement and Learning from Performance Feedback. *Academy of Management Review* 37(2): 211–31. https://doi.org/10.5465/amr.2010.0108.

Kelman, Steven. 2006. Downsizing, Competition, and Organizational Change in Government: Is Necessity the Mother of Invention? *Journal of Policy Analysis and Management* 25(4): 875–95. https://doi.org/10.1002/pam.20212.

Levine, Charles H. 1979. More on Cutback Management: Hard Questions for Hard Times. *Public Administration Review* 39(2): 179–83. https://doi.org/10.2307/3110475.

March, James G. 1991. Exploration and Exploitation in Organizational Learning. *Organization Science* 2(1): 71–87.

March, James G., and Herbert A. Simon. 1993. *Organizations*, 2nd ed. Cambridge, MA: Wiley-Blackwell.

Meyfroodt, Kenn, and Sebastian Desmidt. 2020. Can Rational Planning Stimulate Cooperative Behaviour? How Perceived Self-Efficacy Mediates the Relationship between Strategic Plan Use, Performance Information Use and Strategic Voice by Local Councillors. *Public Management Review*: 1–25.

Moon, M. Jea. 2020. Fighting COVID-19 with Agility, Transparency, and Participation: Wicked Policy Problems and New Governance Challenges. *Public Administration Review* 80(4): 651–6. https://doi.org/10.1111/puar.13214.

Nielsen, Poul Aaes. 2014. Learning from Performance Feedback: Performance Information, Aspiration Levels, and Managerial Priorities. *Public Administration* 92(1): 142–60. https://doi.org/10.1111/padm.12050.

Nielsen, Poul A., and Donald P. Moynihan. 2017. How Do Politicians Attribute Bureaucratic Responsibility for Performance? Negativity Bias and Interest Group Advocacy. *Journal of Public Administration Research and Theory* 27(2): 269–83. https://doi.org/10.1093/jopart/muw060.

Nooteboom, Laura, Eva Mulder, Chris Kuiper, and Robert Vermeiren. 2019. Towards Integrated Care in the Dutch Youth Care System: An Action Based Research. *International Journal of Integrated Care (IJIC)* 19(4): 152.

Ocasio, William. 1997. Towards an Attention-Based View of the Firm. *Strategic Management Journal* 18(S1): 187–206. https://doi.org/10.1002/(SICI)1097-0266(199707)18:1+<187::AID-SMJ936>3.0.CO;2-K.

Posen, Hart E., Thomas Keil, Sangyun Kim, and Felix D. Meissner. 2017. Renewing Research on Problemistic Search—A Review and Research Agenda. *Academy of Management Annals* 12(1): 208–51. https://doi.org/10.5465/annals.2016.0018.

Rap, Stephanie, Denise Verkroost, and Mariëlle Bruning. 2019. Children's Participation in Dutch Youth Care Practice: An Exploratory Study into the Opportunities for Child Participation in Youth Care from professionals' Perspective. *Child Care in Practice* 25(1): 37–50.

Rittel, Horst W.J., and Melvin M. Webber. 1973. Dilemmas in a General Theory of Planning. *Policy Sciences* 4(2): 155–69. https://doi.org/10.1007/BF01405730.

Rutherford, Amanda, and Kenneth J. Meier. 2015. Managerial Goals in a Performance-Driven System: Theory and Empirical Tests in Higher Education. *Public Administration* 93(1): 17–33. https://doi.org/10.1111/padm.12124.

Salge, Torsten O. 2011. A Behavioral Model of Innovative Search: Evidence from Public Hospital Services. *Journal of Public Administration Research and Theory* 21(1): 181–210. https://doi.org/10.1093/jopart/muq017.

Simon, Herbert A. 1965. Administrative Decision-Making. *Public Administration Review* 25(1): 31–7.

———. 1997. *Administrative Behavior*, 4th ed. New York.

Staw, Barry M., Lance E. Sandelands, and Jane E. Dutton. 1981. Threat Rigidity Effects in Organizational Behavior: A Multilevel Analysis. *Administrative Science Quarterly* 26(4): 501–24. https://doi.org/10.2307/2392337.

Van der Voet, Joris. 2019. Organizational Decline and Innovation in Public Organizations: A Contextual Framework of Cutback Management. *Perspectives on Public Management and Governance* 2(2): 139–54.

———. 2021. Policy Preferences in Response to Negative Economic Prospects of COVID-19: A Survey-Experiment among Local Politicians in Four European Countries. *Public Administration*.

Van der Wal, Zeger. 2020. Being a Public Manager in Times of Crisis: The Art of Managing Stakeholders, Political Masters, and Collaborative Networks. *Public Administration Review* 80(5): 759–64. https://doi.org/10.1111/puar.13245.

Visser, Lianne, and Peter M. Kruyen. 2021. Discretion of the Future: Conceptualizing Everyday Acts of Collective Creativity at the Street-Level. *Public Administration Review* 81: 676–90.

Weber, Edward P., and Anne M. Khademian. 2008. Wicked Problems, Knowledge Challenges, and Collaborative Capacity Builders in Network Settings. *Public Administration Review* 68(2): 334–49. https://doi.org/10.1111/j.1540-6210.2007.00866.x.

Wynen, Jan, Jan Boon, Bjorn Kleizen, and Koen Verhoest. 2020. How Multiple Organizational Changes Shape Managerial Support for Innovative Work Behavior: Evidence from the Australian Public Service. *Review of Public Personnel Administration* 40(3): 491–515. https://doi.org/10.1177/0734371X18824388.

Zhu, Ling, and Amanda Rutherford. 2019. Managing the Gaps: How Performance Gaps Shape Managerial Decision Making. *Public Performance & Management Review* 42(5): 1029–61. https://doi.org/10.1080/15309576.2019.1568886.

ZonMW. 2018. "Rapport Eerste evaluatie Jeugdwet—Rapport—Rijksoverheid.nl." Rapport. Ministerie van Algemene Zaken. 2018. https://www.rijksoverheid.nl/documenten/rapporten/2018/01/30/rapport-eerste-evaluatie-jeugdwet [accessed December 18, 2020].

Appendix A: Overview of Policy Solution Categories

In the expert survey, all policy solution categories were presented in a randomized order. The lead-in was: "*Youth care delivery in my municipality can be changed and improved by …*". The creativity ratings by aldermen and youth care managers were given on a scale ranging from 1 to 10.

Solution Category	Frequency	Creativity Rating by Aldermen	Creativity Rating by Directors
1: more collaboration and communication between the municipality and the executive organizations. For example, by consulting professionals and youth care organizations better, and by knowing them as civil servants and municipalities.	105	5.59	5.72
2: more collaboration between youth care organizations. For example, by fusing organizations or by coordinating the supply and working methods of youth care organizations, to avoid double functions.	76	6.14	5.26
3: more collaboration between municipalities in a regional context. For example, by purchasing youth care collectively as much as possible.	81	4.67	5.32
4: to stimulate collaboration by exchanging data between parties involved in youth care through a (national) ICT system. This can be made possible by, for example, relaxing the privacy regulation rules.	25	6.60	6.18
5: more attention to prevention and early identification of problems.	203	6.06	4.92
6: more attention to prevention and early identification among social actors, such as education and youth associations. For example, by supporting them with knowledge in the field of identification and by building in youth facilities.	93	6.59	6.00
7: more attention to prevention and early identification among professional care actors, such as the Child and Family Centre, Municipal Health Services, and the GP. For example, by a greater focus on youth work and accessible parenting help and advice for parents through prevention projects such as the Neighborhood Families project.	123	6.64	6.08
8: To focus more on freely accessible, low-threshold primary youth care offered by local providers such as the General practice-based nurse specialist, school social work, youth and neighborhood teams, or the Child and Family Centre.	208	6.63	5.70
9: to focus more directly on specialized youth care, instead of first deploying care from primary care, that is slowly scaled up. This to prevent too low care efforts and long aftermath.	35	4.34	5.38
10: to receive more financial resources for youth care as a municipality. For example, to be received from the central government (municipal fund) or a contribution from parents (according to their capacity.	143	5.73	4.76
11: a different distribution of financial resources for youth care. For example, by a better distribution of the financial resources among municipalities (as needed) or by disbursing financially within the social domain.	21	5.61	5.98
12: monitoring costs incurred by youth care organizations. For example, by receiving as municipality an indication of the expected costs per client in advance and by steering youth care organizations on timely invoicing of the costs incurred.	53	5.59	3.94
13: controlling costs and stimulating cost awareness of youth care organizations, so that professionals start making decisions with due regard for the cost aspect. For example, by forming financial frameworks in which youth care must be provided (indications with a predetermined maximum amount or a maximum treatment time per client), or by financing based on outcomes.	87	6.15	4.54
14: to steer toward reducing bureaucracy, administration, and overhead costs within municipalities and youth care organizations. For example, by working more based on trust in the (youth care) professional.	192	5.93	6.80
15: adapting market forces in youth care. For example, by abolishing tender within youth care and considering alternatives, or by critically reviewing and improving the tender.	74	6.33	6.50
16: stricter access to youth care. For example, by a stricter demarcation of what is covered by youth care, by having more control as a municipality on needs assessment, or by normalizing behavior and focusing more on the personal responsibility of young people and parents.	349	6.41	4.76
17: facilitate access to youth care. For example, by hiring more youth care professionals, shortening waiting lists, issuing indications more quickly, improving communication about youth care facilities to citizens (social maps), and relaxing the 18–/18+ regulation.	44	4.76	5.36
18: to centralize the organization and financing of (specialized) youth care back to the province or central government and to have the costs for specialized youth care (like before the decentralization) covered by health insurance.	89	4.38	4.68

Solution Category	Frequency	Creativity Rating by Aldermen	Creativity Rating by Directors
19: to decentralize all rules about youth care, so that policy freedom is created at the municipal level and control over youth care remains in the hands of the municipality, instead of regional partnerships or joint arrangements.	39	5.86	4.34
20: offering youth care services as a municipality (instead of hiring external youth care organizations). For example, by offering youth care services as municipality through youth care professionals employed by the municipality.	16	5.67	4.72
21: to generate more specific knowledge about youth care or bringing in knowledge as a municipality (council). For example, by bringing in more youth care expertise or by developing more expertise in analyzing data and contract management of youth care.	25	5.95	5.70
22: to tighten youth care purchasing and contracts. For example, by negotiating rates, purchasing youth care services based on actual needs, and entering multi-year contracts.	66	5.14	4.40
23: to monitor the quality of youth care organizations more. For example, by making result-oriented quality agreements with contracted youth care organizations and by evaluating the effectiveness of their youth care more (e.g. by the youth care inspectorate).	129	5.77	4.36
24: to integrate or coordinate youth care with other policy areas. For example, suitable education and student transport, housing, (free) childcare, poverty reduction, and drug prevention.	34	6.71	7.00
25: a more integrated approach in youth care. For example, by targeting (youth) help to the whole family, appointing one primary therapist and plan per family, organizing access to care (and assessment of needs) from one place, and by offering tailor-made solutions to the request for help.	232	6.38	5.90
26: more research and innovation in youth care. For example, by setting up a large-scale youth care study, including a root cause analysis of common problems, or by working through Evidence-Based Methods and expertise centers.	92	6.42	5.82
27: to stimulate the participation of young people, parents, and citizens in improving the quality of youth care. For example, by appointing a Youth Council or children ombudsman within the municipality and by stimulating citizens' initiatives.	33	6.60	6.40
28: deploy and retain better qualified youth care professionals. For example, by setting up an expertise bank and by recruiting and retaining qualified professionals for the youth care sector through e.g. better working conditions for youth professionals.	34	5.77	5.68
29: to cut back on other domains within the municipality, leaving more money for youth care.	7	3.53	4.72
30: To be more selective with the contracting of youth care organizations (e.g. on quality) by adapting the unlimited entry of youth care organizations in a municipality (the "open house construction").	82	5.59	5.34
31: giving youngsters and parents more responsibility and control within the youth care process. For example, by letting the parents (and young people) participate more actively in the youth care process.	22	6.34	6.22
Total	2,812		

Appendix B: Categorization Process of Ideas Generated in Survey Experiment

Step 1: Scanning the main content of the ideas and formation of the initial set of categories.
The generated ideas were listed in alphabetic order in a text document, scanned and organized into frequently named subjects. At first, the following initial categories were identified:

- collaboration;
- prevention and early detection;
- money, financial resources;
- centralizing, undoing the decentralization;
- the role of parents;
- bureaucracy, administration, overhead (costs);
- market dynamics;
- procurement, contracting of care providers;
- the role of the general practitioner;
- improved evaluation of services.

Step 2: Assigning ideas to specific categories and reduction of categories.
After the first round of categorizing, the policy solutions were categorized into 100 different, more specific categories. Reviewing the content of the categories, several categories were merged to achieve reduction of 50% in categories. After another round of reduction, the total amount was reduced to 28 categories. Some ideas were not assigned to the 28 categories (see Appendix A). The following decision-rules were applied to exclude responses:

- The solution was too vague, unintelligible, or unrelated to the topic at hand;
- A (rhetorical question) was provided, rather than a solution;
- A problem diagnosis was provided rather than a solution;
- The response was too municipality-specific to be categorized;
- The response contained to alternative course of action.

Step 3: Control round of categorizing and pilot testing.
After all solutions were assigned to a category, the assignment of solutions to categories was checked by the researchers for accuracy. In this process, some ideas were reassigned to a category that fitted the solution better. After getting feedback on the trial version of the expert survey by one alderman and one youth care manager, three categories were split up into two distinct categories, resulting in solution categories 29, 30 and 31 (see Appendix A).

Appendix C: Flowchart of the Procedures of the Survey Experiment