



AARHUS UNIVERSITY



Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Andersen, S. C., & Moynihan, D. P. (2018). How Do Socially Distinctive Newcomers Fare? Evidence from a Field Experiment. *Public Administration Review*, 78(6), 874-882.

<https://doi.org/10.1111/puar.12957>

Publication metadata

Title:	How Do Socially Distinctive Newcomers Fare? Evidence from a Field Experiment
Author(s):	Simon Calmar Andersen, Donald P. Moynihan
Journal:	<i>Public Administration Review</i> , 78(6), 874-882
DOI/Link:	https://doi.org/10.1111/puar.12957
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

How do Socially Distinctive Newcomers Fare? Evidence from a Field Experiment

Simon Calmar Andersen, Department of Political Science, and TrygFonden's Centre for Child
Research, Aarhus University

Donald P. Moynihan, McCourt School of Public Policy, Georgetown University

Biographies

Simon Calmar Andersen is Professor at the Department of Political Science and Director of TrygFonden's Centre for Child Research at Aarhus University. He combines field experiments, surveys and administrative data to examine different aspects of public policies and management strategies and their impact on organizational performance, especially within education. He has published work in Proceedings of the National Academy of Sciences, the Journal of Public Administration Research and Theory, and Public Administration Review among others.

Donald P. Moynihan is the inaugural McCourt Chair at the McCourt School of Public Policy at Georgetown University. He studies the performance of public organizations, and administrative burdens that citizens experience in their interactions with public services. With Pamela Herd, he is the co-author of *Burdened: Policymaking by Other Means* available from the Russell Sage Foundation.

Abstract

New hires offer a mixed blessing. They can spur teams to reflect on their processes in ways that encourage learning. But organizational newcomers may also struggle to achieve inclusion. We examine how newcomer's experiences in public organizations depends upon their social distinctiveness. While diversity is usually framed in terms of biodemographic factors such as race and gender, educational background offers another form of social distinction. Educational differences may trigger psychological responses such as negative social categorization, as well as serve as an observable criterion by which professional status and power is allocated. Using a field experiment, schools were provided two types of new hires: those who shared the educational background of existing teams, and more socially distinctive newcomers. Both types of newcomers led to heightened team reflection processes relative to a control group. However, oldtimers were less accepting of educationally distinct newcomers, viewing them as less competent and cooperative.

Evidence for Practice

- Workplace diversity is most frequently described in bio-demographic terms such as race and gender. Another relevant way to think about diversity is in terms of the different perspectives people bring to their jobs. This kind of functional diversity matters to workplace outcomes, and is shaped by factors such as educational background, experiences, and skills.
- In a field experiment, new hires increased team reflection on staff work and goals.
- Newcomers who lacked the dominant educational background of organizational oldtimers are viewed as less competent relative to newcomers who are less functionally diverse.
- Public managers can use functional diversity to improve learning within organizations, while guarding against the risk of conflict and negative social categorization that socially distinctive newcomers face.
- Communicating the benefits of functional diversity may help employees overcome biases against socially distinctive newcomers.

Integrating new hires into existing teams is a challenge. From an organizational perspective, newcomers offer mixed blessings. They can disrupt old processes, generate team reflection, and improve outcomes (Chen, 2005), even as they struggle to gain acceptance (Rink et al., 2013). One troubling possibility is that those defined as members of an out-group fail to achieve full inclusion. Newcomers with different backgrounds may be especially vulnerable to exclusion in public organizations where professional training is tied to status (Freidson, 2001). A more optimistic possibility is that newcomer's distinctive capacities are valued, leading to both acceptance and innovation (Foldy 2004). Understanding how newcomers can serve as an asset rather than a source of conflict is a challenge that every public manager faces.

Public administration research on diversity has predominantly focused on bio-demographic factors such as race, ethnicity, and gender (Sabharwal, Levine and D'Agostino 2016). However, newcomers possess a multitude of identities, which may be more or less salient in different intergroup relations (Tajfel and Turner, 1979; Zhu, Shen and Hillman, 2014). Social characteristics such as educational background may serve an unappreciated role in public organizations, where unions may see the introduction of employees with different background as a threat to collective representation.

We utilize a randomized controlled field experiment to test how the educational background of newcomers matters to team receptivity, i.e. "the ways in which small groups and work teams respond to newcomers" (Rink et al 2013). Forms of social distinctiveness like educational background may be a double-edged sword (Joshi and Roh, 2009). Newcomers may make a positive difference to the organization, as their mere presence may cause greater reflection on team processes that can ultimately improve outcomes. We therefore propose that newcomers will stimulate more team reflection than teams without newcomers, and that more

socially distinct newcomers will stimulate more team reflection than socially similar newcomers. A newcomer defined as an out-group member may face more negative social categorization than a newcomer that shares characteristics with oldtimers. We therefore also propose that educationally distinct newcomers face more conflict and will be less accepted by existing team members.

Our article offers a number of contributions. First, we import theoretical concepts from organization studies that can offer new insights to salient issues for public organizations – onboarding new hires, diversity, organizational learning and conflict. Using these concepts to better understand the real-world dynamics that socially distinctive newcomers face has direct and practical implications for managers. This is especially true for organizations with a stated goal of managing for inclusion, as is common in the public sector (Quick and Feldman 2011; Oberfield 2014). Second, we integrate newcomer theory with research on team diversity, and draw from organization studies to adopt a functional perspective on diversity, examining how the distinctiveness of the educational background of the newcomer matters. Third, while prior work offers mixed evidence that socially distinct newcomers face higher friction as a result of negative social categorization (Rink et al. 2013), we offer evidence of such an effect in a field experiment that allows us to address endogeneity and selection issues that arise in observational studies.

The experiment added new employees to create work teams in Danish schools. The treatments were systematically varied by educational background – a standard and easily observable indicator of functional diversity. Some hires shared the educational background of existing teachers, and some did not. We find that existing staff are less accepting of educationally distinct newcomers, evaluating them as less cooperative and competent relative to new employees that share their educational background. These negative social categorizations

occur despite objective evidence that socially distinct new hires performed just as well as other new hires. At the same time, both types of newcomers were associated with similar levels of team reflection processes, which were significantly higher than for a control group that received no new employees. The results therefore show that educational background is a salient form of social distinction resulting in more distinctive newcomers struggling to gain inclusion, even as they generate the same types of benefits as newcomers that share the educational background of existing team members.

The next sections integrate the common threads between newcomer and diversity research to motivate our hypotheses. We then present our data and methods, and results, before concluding by considering lessons for practice.

Newcomers as Disruptive Influences

Within organization studies, newcomers are team members who had not been part of the group at some previous point in time, and are often characterized as disrupting group norms, communication processes, or social functioning (Phillips, Liljenquist, and Neale 2009). Teams tend to prefer the familiar, and therefore are generally suspicious of outsiders (Foldy 2004; Moreland and Levine 2006; Joardar et al., 2007), resulting in stultified work processes and suboptimal performance (Katz 1982). Resistance to competing perspectives may be especially pronounced for public organizations, where the dominance of professions, structural protections, and the influence of unions can serve to buffer status quo arrangements from new types of employees.

We draw from Rink et al.'s (2013) review of newcomer research, which argues that how teams receive newcomers has been understudied relative to studies of how newcomers adapt, or

are socialized. Within public administration research, attention to team receptivity is most visible in research on collaboration between different organizations. In such settings, heterogeneity in backgrounds is portrayed both a motivation to collaborate, and a limit of group efficacy (Ospina and Saz-Carranza 2010; Leach et al 2010). Other work connects the participation of women and ethnic minorities with an inclusive work climate (Andrews and Ashworth 2015), while also stressing the importance of diversity management programs (Choi and Rainey 2009).

One measure of how teams function is “team reflection,” i.e., active consideration of work processes (Rink et al. 2011). In the field that we study, education, the topic of team reflection is of interest in its own right, though often framed as teacher “collaboration” or “coordination” (Pounder 1998). Such team efforts to discuss work processes is associated with better student outcomes (Goddard, Goddard and Schannen-Moran 2007).

Newcomers do not have to actively contribute or be solicited for knowledge to generate team reflection. Indeed “team knowledge utilization” is a separate standard for examining if teams are actively drawing on the skills, attitudes and capacities of newcomers. The team reflection category points out that the mere disruption introduced by a newcomer in a team is enough to start reflective processes. We therefore propose that newcomers will generate greater team reflection processes relative to comparable settings where no newcomers are introduced. Both research on newcomers and diversity generally support this proposition (Bell et al., 2011; Joshi and Rohi, 2009; Phillips, Liljenquist and Neale 2009; Rink et al. 2011), though without strong field experiment evidence.

H1: The introduction of a newcomer will result in greater team reflection processes.

Combining Social Distinctiveness and Newcomer Status

This section considers how socially distinct newcomers may influence team reflection processes more so than newcomers that are more similar to oldtimers. The role of social distinctiveness for newcomers has not been widely studied, with mixed evidence on how it matters (Rink et al. 2011). One reason for the lack of clear evidence is, according to Phillips, Liljenquist and Neale (2009), a tendency to conflate newcomer status and social distinctiveness. They use an experimental design to distinguish these factors on team processes and ultimately performance. Tasking college students to solve a murder mystery puzzle, the design ingeniously used membership of fraternities and sororities to create variation in the social distinctiveness of new team members. Teams with a socially distinct member saw greater resistance to including the new member, but also greater team reflection and task success. Even if socially distinct newcomers do not articulate new perspectives, the way in which they approach their job may differ from current employees, perhaps compelling other employees to adjust their own approach. The presence of socially distinct actors may also cause other team members to prepare more before meetings (Loyd et al., 2012).

Studies of diversity in organizations also examine how out-group members are accepted, and the effect of out-group members on team processes. Beyond newness, what makes newcomers socially distinctive in practice will be tied to some aspect of diversity, arising “from any attribute people use to tell themselves that another person is different” (Williams and Reilly, 1998; 81). These could be bio-demographic factors such as age, race, gender, or functional diversity factors that alter how people encode and solve problems (Long and Page, 2004; 1685), such as educational background, experiences, skills and perspectives (Joshi and Roh, 2009).

In this article we focus on the social distinctiveness of new employees in terms of the functional diversity that arises from a different educational background. The literature on

bureaucratic politics stresses how different educational background matters for public organizations: for example, the increased influence of economists and engineers over lawyers in regulatory agencies altered how work was approached, tasks defined, and ultimately what organizational outcomes were achieved (Eisner and Meier 1990; Wilson 1989). In prior work, we have shown that how public managers respond to educational diversity is contingent on their organizational culture (Andersen and Moynihan 2016), but such research is the exception within public administration scholarship, which has largely excluded attention to educational background and other forms of functional diversity (Choi and Rainey 2009; Oberfield 2014; Sabharwal 2014, 779).¹

In part, the focus on bio-demographic factors in public administration scholarship may reflect legal requirements arising from the public setting, and may be partly due to a normative interest in equity and representative bureaucracy in the service of democratic principles (Sabharwal, Levine, and D'Agostino 2016). As a result, other conceptions of diversity or social distinctiveness that feature in the broader field of organization studies are largely absent in public administration. Sabharwal, Levine, and D'Agostino (2016, 15) conclude their analysis of 75 years of public administration diversity research by calling for alternative perspectives: "Apart from examining race and gender, future researchers can adopt a more synergistic view of diversity that values diversity in all its forms."

To learn more about the effect of social distinctiveness therefore demands looking beyond public administration to the broader field of organization studies. Here, research offers support for both positive and negative accounts of social distinctiveness (Bell et al., 2011; Joshi and Rohi, 2009; Horowitz and Horowitz, 2007; Mannix and Neale, 2005; Williams and Reilly, 1998). On one hand, different meta-analyses find that bio-demographic diversity is associated

with either no effect (Horowitz and Horowitz, 2007) or small negative effects on team processes and performance (Bell et al., 2011; Joshi and Rohi, 2009). On the other hand, measures of functional diversity are associated with positive effects. Joshi and Roh (2009) find that functional diversity has a more substantive positive effect on measures of team performance than bio-demographic measures have a negative effect. Another meta-analysis found that functional diversity had a positive impact on team decision-making, creativity, innovation, and problem solving (Horowitz and Horowitz, 2007). Bell et al.'s (2011) meta-analysis also finds that variety in functional background and educational background is associated with greater team creativity and innovation. Diversity in functional background also has a small positive effect on measures of performance, and educational diversity is associated with higher performance for top-level managerial teams (Bell et al., 2011).

More positive empirical accounts of diversity emerge from laboratory experiments that take an information/decision-making approach (Mannix and Neale, 2005). Here, the causal mechanism, which informs our next hypothesis, is that diverse groups outperform homogenous groups by benefiting from a wider array of information, experience, skills, and perspectives. These resources, as well as the processes of debate and knowledge exchange, foster creativity and problem solving, and prevent group-think and myopia. For example, when students in a lab experiment were told they would meet with and discuss policy issues with students with different political ideologies, they tended to prepare more, suggesting that social distinctiveness primes greater pre-meeting elaboration (Loyd et al. 2012). Some field studies provide additional support for the information/decision-making approach in non-laboratory settings, tying organizational-level diversity to positive outcomes (Feng and Nagar, 2013; Østergaard et al., 2014). Formal models of diversity also rely on assumptions derived from an information/decision-making

approach to posit positive benefits of diversity. Long and Page (2001, 2004) demonstrate that a random selection of diverse agents outperforms less diverse teams of higher-performing individuals. “This result relies on the intuition that, as the initial pool of problem solvers becomes large, the best-performing agents necessarily become similar in the space of problem solvers. Their relatively greater ability is more than offset by their lack of problem-solving diversity” (Long and Page, 2004; 1685).

Consistent with the information/decision-making approach to diversity we propose that more distinct newcomers generate greater team reflection processes.

H2: Educationally distinct newcomers will generate greater team reflection processes than educationally similar newcomers.

Newcomer Acceptance

In this section we examine another form of team receptivity identified by Rink et al. (2011): “newcomer acceptance,” where the newcomer is fully included as a full team member. Newcomer acceptance can be assessed in a variety of ways, including oldtimer attitudes toward newcomers, their assessments of newcomer attributes, perceptions of whether newcomers are helping or hurting team dynamics, or newcomer’s own assessment of whether they feel accepted.

Prior research suggests that more socially distinct newcomers face lower acceptance than more similar newcomers (Joardar et al., 2007; Schwierer and Glunk 2008). Such resistance is explained by two broad theoretical frameworks: social categorization and similarity-attraction models (Williams and O’ Reilly, 1998; and Mannix and Neale, 2005). Social categorization is the process by which observable cues are used to differentiate between individuals, classifying others in terms of value, ability and whether they belong in an in-group or out-group. A

judgment that an individual does not belong to the dominant group generally results in negative associations (Tajfel and Turner, 1979). Even if the basis of differentiation is objectively a poor predictor of an individual's ability to complete a task, social categorizations tend to correlate with negative processes and outcomes, by triggering biases, increasing conflicts between individuals, and failing to include the perspectives of the negatively categorized individuals.

Similarity-attraction theory mirrors the logic of social-categorization, proposing that individuals with common attributes enjoy smoother working relationships and better outcomes. As a result, diversity will lead to greater conflict and lower performance. Social categorization and similarity-attraction theory rely on a common causal mechanism that becomes the basis for our third hypothesis: individuals are generally negatively disposed toward those who are different from them, and use bio-demographic factors such as age, gender, or race offer easily-observable cues that lend themselves to judgments, often informed by prejudices.

Educational distinctiveness might be expected to be less associated with the type of negative social categorization that previous work has found with bio-demographic diversity. However, professionals may be especially sensitive to educational qualifications, and try to protect their privileges by excluding workers with distinct educational backgrounds (Freidson, 2001). In such circumstances, more educationally distinct newcomers might experience more resistance from oldtimers skeptical of other forms of educational training, even as shared educational backgrounds allows oldtimers to look past other newcomer differences (Zhu, Shen and Hillman, 2014). Our final hypothesis therefore draws from the similarity attraction and social categorization models to propose that educationally distinct newcomers will receive less acceptance.

H3: Educationally distinct newcomers are less accepted than educationally similar newcomers.

Data and Methods

We exploit a randomized controlled field experiment in Denmark conducted in cooperation with and funding from the Danish Ministry of Education, who were interested in new options for co-teaching in schools.² In Denmark, schools are operated by 98 multi-purpose municipalities, and all were invited to participate in the experiment. Among the 68 applicants, 12 municipalities were selected, with purposeful sampling to ensure municipalities varied in terms of geography and school size. The 68 municipalities that applied were similar to the 30 that did not apply on a number of observable characteristics ranging from number of inhabitants over expenditure on public schools to share of students with immigrant background or low-educated parents. The only statistically significant differences relate to number of schools and geographical size of the municipality. Details of the randomization procedure as well as a test showing that the experimental groups balance on observed variables are presented in the appendix (see Figure A1 and Table A1). The self-selection of municipalities and schools into the application procedure implies that generalization of results to all schools should be done cautiously. However, since schools that were enrolled in the trial were randomly assigned to the experimental conditions, the estimates of the intention-to-treat effects should be unbiased for the participating schools.

Rink et al.'s (2013) comprehensive review of newcomer research describes a mix of experimental or field-based approaches, but few combinations of these approaches. While there is some research on the role of social distinctiveness to team receptivity, buttressed by studies of

diversity effects, the experimental design offers a number of advantages over existing work. First, it has the advantage of avoiding the problem that Phillips, Liljenquist and Neale (2009) identify in newcomer research: the tendency to conflate newcomer status with being categorized as a member of an out-group. While newness is one way that employees categorize each other, diversity research points to other factors that might result in two newcomers experiencing different levels of team receptivity. Our experimental design allows us to discern how much educational distinctiveness matters *among* newcomers as well as relative to teams without newcomers.

Our experiment is also grounded in an actual work environment. Laboratory experiments, by design, cannot model the experience of an organizational context. Another criticism of laboratory experiments in this area is that the groups created are artificial, deal with short-term tasks, and lack the ongoing social interactions of actual groups dealing with diversity (Williams and O' Reilly, 1998). While some experimental work features a longer-term set of tasks for groups to solve, such research tends to study students completing the exercise as part of their classwork (Polzer, Milton and Swarm, 2002; Watson, Jumar and Michaelson, 1993; Van der Vegt and Van de Vliert, 2005). Such studies may be less suited to examine what social characteristics are salient in real organizations. We examine the effect of newcomers on team receptivity 7-8 months after the introduction of the newcomer, a longer time-frame than is typical for laboratory experiments.

Observational studies also cannot address selection processes. In the context of our research question, this is a salient consideration because of the very real risk that socially distinct potential employees will not be selected into the organization, or that selection is based on unobserved factors. For example, Zhu, Shen and Hillman (2014) show that even when

demographically diverse actors are selected into the organization, this is often on the basis of some other shared characteristic that allows the newcomer to be categorized into the in-group. In observational studies, the nature of organizational diversity may therefore be endogenous with measures of outcomes, limiting any claims on causality.

The newcomer experience is ideally suited to a longitudinal approach that can track experiences (e.g., Cable, Gino and Staats, 2013), combined with a field experiment design that places the research in a real political context while addressing endogeneity problems. The next section explains how we apply such an approach.

Treatments: Newcomers and Social Distinctiveness

The field experiment gave participating schools resources to hire different types of new human resources to create work teams. How those employees differ from existing employees creates variation both in terms of social distinctiveness, allowing us to examine how variation in the introduction of these factors is associated with newcomer acceptance and team reflection processes.

The use of such assistants is itself salient as a policy issue. In many countries, teacher assistants constitute a substantial share of the workforce, and the decline of teacher unions paves the way for the use of non-traditional teachers. In Iceland the teacher–assistant ratio is 4:3 (OECD 2014). In England, the number of teaching assistants has tripled since 1997, and in 2011 they comprised 24% of the mainstream school workforce (Blatchford et al. 2012). In Denmark at the time of the field experiment, only educated teachers were allowed to be responsible for the ordinary teaching (with few exceptions, the Public School Act (“Folkeskoleloven”) 2010, §28). Teacher assistants were allowed, but might have been seen as competing for jobs that teachers

with a traditional educational background could take. A school reform enacted a few years after the field experiment encouraged schools to integrate child and youth workers and kindergarten teachers in schools' teaching staff. The use of teaching assistants and co-teachers in teacher teams is therefore politically salient and highly relevant for school managers in many countries.

In the field experiment, schools were randomly assigned to a control group or one of two treatment conditions:

1. **Control group:** School principals in the control group did not receive any additional human resources. There is therefore no additional diversity added to the workforce.
2. **Educationally similar newcomer:** Schools were granted a budget of approximately \$25,000 to hire new employees, but with the stipulation that the hires have a BA in teaching. Newcomers share the same educational background as existing teachers.
3. **Educationally distinctive newcomer:** Schools were granted a budget of approximately \$25,000 to hire new employees, with the stipulation that these hires have any educational background other than a BA in teaching. These new employees are therefore socially distinct from existing employees in terms of their educational background. In practice, these hires have been used as co-teachers, putting them in the same task interdependent as the educationally similar newcomers.

All schools, including the control group, delivered data to the study. Hence, all were aware that they were part of an experiment. Therefore, the potential that the findings are a function of a Hawthorne effect (i.e. the behavior changes as a function of being studied) is reduced.³ Further, both treatment groups received a treatment, so any differences between treatment groups cannot be caused simply by receiving any treatment.

Dependent Variable: Newcomer Acceptance

To measure if newcomer's gain inclusion or face conflict, we compared existing employee perceptions about the new hires in the two treatment groups. These measures were collected via an employee survey provided 7 to 8 months after the introduction of the treatments, with a response rate of 63% of a target group of 787 employees (two teachers per classroom and a little less than one fourth in each treatment group and little more than half in the control group, see table A1 in the appendix). Since our goal is to understand how existing organizational staff responded to the addition of new employees, these new employees are not included among the respondents. The respondents were instead the oldtimer team members who were working with newcomers.

We define newcomer acceptance using two of the standard ways that Rink et al., (2012) describe: the categorization of newcomer attributes and the perceived effect on team dynamics. Our first measure is an indicator of perceived newcomer competence. The second reflects the perceived effect of the newcomer on team dynamics by assessing their perceived cooperativeness.

To measure competencies of the new team members, an index made up of six items (Cronbach's alpha = .95) was assembled: "To what extent could the extra person handle the assigned functions?; The professional competencies of the additional persons are satisfactory; The personal competencies of the additional persons are satisfactory; The pedagogical competencies of the additional persons are satisfactory; The competencies of the additional persons in relation to students' wellbeing and social relations are satisfactory;" and one reverse-coded item: "To what extent do you agree or disagree that the following factors have been a

barrier to successful implementation of the trial at your school? The additional person lacked the necessary competencies.”

Cooperation was measured by two items (Cronbach’s alpha = .90), “To what extent are you satisfied with the cooperation with the additional person” and “To what extent do you agree that the additional person has built good relations to other teachers.” For both sets of questions, response rates were on a five point Likert scale ranging from (1) a very low extent to (5) a very high extent. The items for each measure were summarized in two indexes based on their factor scores in principal component factor analyses. The scales were standardized with a mean of 0 and a standard deviation of 1.

Dependent Variable: Team Reflection

Our assessment of existing staff is consistent with what Rink et al. (2013) categorize as team reflection processes, while grounded in the particular function of education that we study. We treat team reflection as efforts to examine teaching and helping students learn rather than resting on a more abstract conceptualization of reflection unmoored to actual tasks. In education research, this is sometimes framed as “collaboration” (Goddard, Goddard and Schannen-Moran 2007).

In the field of education there has been, in Denmark and globally, growing pressure on teachers to improve performance (Hvidman and Andersen, 2014). Doing so requires understanding what it is that teachers do and how that relates to organizational goals. For example, techniques such as value-added modeling (Harris, 2011), or monitoring every aspect of what teachers do in a classroom (Thomas and Staiger 2012) seek to identify both which teachers are effective and what makes for an effective teacher. Both elaboration processes that seek to

understand student needs and progress and engaging in peer discussion about teaching techniques can therefore represent highly salient forms of problem-solving to the organizations we study.

Teachers were asked to respond to the following questions “to what extent do you do the following activities?: “Discuss the development of specific students’ learning with other teachers or other employees at the school.”; “Evaluate my teaching with other teachers or other employees at the school”; “Observe the teaching of other teachers and give feedback.” Again, response scale ranged from 1 (a very low extent) to 5 (a very high extent), and loadings from a principal component factor analysis were used to summarize items into a standardized index with a mean of 0 and a standard deviation of 1. Cronbach’s alpha for these three items was .69.

Results and Discussion

Table 1 provides the results for the hypotheses related to team reflection processes, while table 2 provides the results related to newcomer acceptance. Since the Cronbach’s alpha for our items on team reflection is near the margins of acceptability, we ran separate models for the three components of the index (model 1-3) as well as for the overall index (model 4). Because these items ask about patterns of team reflection, we include the control group who received no new employees as the reference category.

Collectively the results show that the introduction of newcomers, whether they share the dominant educational background of existing staff or not, are associated with more team reflection processes. This offers support for our first hypothesis. Our second hypothesis tested if more socially distinct newcomers generated more team reflection processes. Here, the results do not support the hypothesis. If anything, the results suggest that more socially similar newcomers

are associated with greater team reflection, although the difference between the treatments is not significant for any coefficient. The coefficient for the overall index for socially similar newcomers is higher, though again not statistically different from socially distinct newcomers. Since the dependent variable is standardized, the coefficients can be interpreted as the effect of the treatments measured in standard deviations. Introducing newcomers thus increase team reflection with between .32 and .46 of a standard deviation. Since we do not know of similar field experimental studies, assessing these effect sizes is difficult, but they seem to be non-trivial.

Table 2 addresses our third hypothesis, that socially distinct newcomers struggle to achieve full inclusion. Since the control group does not have any newcomers to evaluate they cannot be included in this analysis. We therefore compare the acceptance of socially similar and socially distinct newcomers. The results are consistent with the hypothesis. Newcomers that did not share the educational background of existing teachers were significantly more likely to face negative social categorization, being viewed as both less competent and less cooperative. Since the dependent variables are standardized the coefficients mean that educationally diverse newcomers are perceived .70 of a standard deviation less competent and .47 of standard deviation less cooperative. Again, we see these as non-trivial effect sizes, but do not have other studies to compare against.

One possible explanation for these effects is that processes of social categorization are indeed justified, that those who come from a different educational background are simply less able to play a useful organizational role in schools relative to those who are trained to be teachers. However, there are a variety of reasons to doubt this explanation. While this explanation might account for differences in perceived competence, they do not explain why those with a different educational background would be systematically viewed as less

cooperative. It is also the case that those selected without a background in teaching were hired via a more open selection process that should generate a more capable pool of candidates, and many had qualifications to work with children.⁴ Finally, and most persuasively, a post-experiment evaluation of the effects of the treatments found that while the co-teachers had a significant impact on improving student outcomes in some dimensions (consistent with hypothesis 1), there was no systematic difference in performance between co-teachers who shared the educational background of teachers and those who did not (Andersen et al. 2018). Overall then, the balance of evidence does not support the explanation that the oldtimers resist socially distinct newcomers because they are less competent. Instead, the evidence is more consistent with the possibility that negative social categorization arises from underlying prejudices of a dominant group towards new colleagues who came from a different background. Such prejudices may be confounded in a public sector setting where members of well-established professions view educational credentials as de facto indicators of competence, and see out-group newcomers as competing for the same jobs. The disruption that newcomers bring by their presence may be perceived as a threat to existing status, influence, and job security (Ospina 1996; Eisner and Meier 1990).

The results provide experimental evidence for the doubled-edged nature of the newcomer experience, but only for those who are socially distinctive. In some respects, these findings are consistent with prior research on diversity. While all of the treatments introduced newcomers, socially distinct newcomers were associated both with team reflection processes but also lower acceptance by oldtimers. The experimental design helps to exclude alternative explanations for the source of these assessments. It is variation on educational background that drives the negative assessments of competence and cooperativeness.

The framing of diversity as a double-edged sword, and the empirical evidence that supports this framing, tempts a small logical leap to propose that the positive and negative aspects of diversity are inextricably tied together: the conflict and friction is a necessary step in generating creativity and problem-solving (e.g. Pelled, Eisenhardt and Xin, 1999). While this may sometimes be the case, the results of the experimental treatments on team reflection processes are interesting partly because they refute the idea that the positive aspects of diversity can only be gained by going through experiences of conflict. New teachers who shared the educational background of existing staff generated team reflection processes without experiencing negative social categorization. This finding is consistent with Rink et al.'s (2013) call for caution when introducing socially distinct newcomers. On the other hand, we show that socially distinct newcomers can foster team reflection processes even if they are not fully accepted as team members. This suggests that the benefits of newcomers may not depend on their full inclusion.

Conclusion

When organizations commit to the notion of diversity, this creates an obligation to follow-through on that commitment (Choi and Rainey 2010). Otherwise, employees become disillusioned when promises are not kept. Such commitments, in turn, create a need for employers to understand the different aspects by which diversity matters in the workplace. While public administration scholarship considers bio-demographic factors, other forms of social distinctiveness may be just as salient to how employees judge one another and for work outcomes.

Our analysis sheds causal insights on functional diversity via a field experiment that tracks the introduction of socially distinct newcomers into organizations. The key finding is that

newcomers generate positive team reflection processes, but that socially distinctive newcomers face negative social categorization. Because we focus on schools, one obvious limitation is generalizing to other types of organizations. However, the field of education is important in its own right, and has been experiencing many of the types of changes we include in our research, such as closer observations of teacher performance in the classroom, and more experimentation with the use of teaching assistants, as well as a general interest in team reflection processes as a form of collaboration (Pounder 1998).

Another limitation is that we do not observe the outcomes over a more sustained time period. Time matters, since newcomers eventually become oldtimers as their relationship with their organization evolves (Solingers et al., 2013). Socially distinct newcomers might benefit especially from being seen to commit to the organization over a longer time period (Rink and Ellemer, 2009; Pelled, Eisenhardt and Xin, 1999; Schippers et al., 2003). Intergroup beliefs and behavior take time to change (Pettigrew 1998), and more recent research emphasizes the dynamic nature of team development (Srikanth, Harvey and Peterson, 2016). Ultimately, the possibility that time matters to diversity effects does not negate the design we employ, but instead points to the underlying need for research that further accounts for temporal dimensions in field settings.

Public organizations seeking to expand the functional diversity of its workforce need to understand and respond to the particular types of social distinctiveness that are at risk of triggering negative social categorizations. This may especially be the case in the political context of public organizations where educationally distinctive newcomers may evoke more deep-rooted professional opposition. Previous work has pointed to ways in which education was sometimes used as a form of capital to attain status and power (Ospina 1996). Education is just one aspect of

functional diversity, and this article demonstrates its salience in a public setting. We also offer evidence that in a professional setting that values certain types of educational credentials, those that lack those credentials may face negative social categorization. For those concerned with inclusive public management (Feldman and Khademian 2000; Quick and Feldman 2011), we identify another type of exclusion that public sector bureaucracies perform, which is to resist the inclusion of non-traditional types of employees.

For public managers overseeing teams, awareness of the double-edged nature of newcomer social distinctiveness represents both a challenge and opportunity. Judiciously adding newcomers to teams can spark constructive reflection and dialogue about how team processes work in the first place. How might resistance to functional diversity be overcome? Structural factors, such as civil service rules, can weaken the potential for education-based or professional biases in the workplace (Ospina 1996). Managers also have to watch for how socially distinctive newcomers are treated, and take proactive steps to communicate the value of the background of the socially distinctive newcomer, the value of functional diversity more generally, and be ready to intervene directly when they perceive newcomers are treated as second-class organizational citizens.

The results raise – but by no means conclude – the question of how the positive and negative aspects of the newcomer experience are bound together. Newcomers appear to generate closer attention to work processes. Employees who are embedded in teams and work processes cannot be easily ignored, increasing the potential for conflict and negative social categorization, but also forcing team reflection behavior to occur if tasks are to be achieved. Future work on this topic could study the experience of such employees as an alternative means to foster inclusiveness. More broadly, given the limited existing attention to newcomer status and

functional diversity in public administration, there is ample opportunity to incorporate these concepts into the study of diversity in public organizations.

Endnotes

1. One representation of strong biodemographic of diversity research in public administration is exemplified by the codes Sabharwal et al (2016, 7) used to make sense of 75 years of public administration diversity research: “The various dimensions of diversity that were coded in this study include race (ethnicity/minority), gender, AA/EEO, representative bureaucracy, age, disability, religion, socioeconomic status (SES), sexual orientation, immigration, generations, veterans, diversity management, cultural competency (including culture), and other. ” Of these, 63% of articles focused on race, 55% on gender, 19% on affirmative action/equal opportunity, 21% on age, and 18% on representative bureaucracy (articles could include multiple topics, so totals do not add up to 100%). No other topic was featured in more than 9% of articles.

2. The trial has been used in previous studies to examine effects of co-teaching on student learning (Andersen et al 2018) and manager’s reactions to the recruitment of new hires (Andersen and Moynihan 2016). In this study we focus on existing teachers’ reactions to the introduction of new team members. We therefore leave out data on schools that in a separate and unrelated treatment condition were assigned teacher supervisors, which were not part of the teacher teams.

3. In a narrow definition, Hawthorne effects arise when individuals modify their behaviour in response to their awareness of being observed as part of a trial. This may be caused by the process accompanying experimentation such as “signing of release forms, reading instructions etc.” (Levitt and List 2011, 227). This form of Hawthorne effects can be controlled for if the control group sign the same forms and take part in the same data collection activities as the treatment group (Levitt and List 2011, see also Adair, Sharpe and Huynh, 1989). In this experiment, all schools and teachers (i.e. both treatment and control group) were informed prior to the intervention that they were part of the research project in which they would be observed and required to contribute to the data collection. Effects of being observed should therefore affect both treatment and control group. Hawthorne effects may also arise through more subtle channels, though. For instance, if the presence of the newcomers repeatedly reminds teachers in the treatment condition that they are part of a trial and thereby make them react in ways they would not do outside of the trial. However, if teachers react to the aides by increasing their own effort this may be an effect related to the treatment, which would most likely also generalize to situations outside the trial.

4. The assistants without a BA in teaching had wide variety of educational backgrounds. The main categories were child and youth worker (pedagogue): 29%; Master degree (University): 16 %; Teacher BA-student: 9 %; Tradesman: 4 %; Other: 43 %.

References

- Adair, John G., Donald Sharpe and Cam-Loi Huynh. 1989. Hawthorne Control Procedures in Educational Experiments: A Reconsideration of Their Use and Effectiveness. *Review of Educational Research* 59(2): 215-228.
- Andersen, Simon Calmar and Donald P. Moynihan. 2016. How leaders respond to diversity: The moderating role of organizational culture on performance information use. *Journal of Public Administration Research and Theory* 26(3): 448-460.
- Andersen, Simon Calmar and Louise Beuchert and Helena Skyt Nielsen and Mette Kjærgaard Thomsen. 2018. The Effect of Teacher's Aides in the Classroom: Evidence from a Randomized Trial (January 29, 2018). Available at SSRN: <http://dx.doi.org/10.2139/ssrn.2626677>
- Andrews, Rhys, and Rachel Ashworth. 2015. Representation and Inclusion in Public Organizations: Evidence from the U.K. Civil Service. *Public Administration Review* 75(2): 279–88.
- Bell, Suzanne T., Anton J. Villado, Marc A. Lukasik, Larisa Belau, and Andrea L. Briggs. 2011. Getting Specific about Demographic Diversity Variable and Team Performance Relationships: A Meta-Analysis. *Journal of Management* 37(3): 709-743.
- Blatchford, Peter, Anthony Russell, and Rob Webster. 2012. *Reassessing the Impact of Teaching Assistants: How Research Challenges Practice and Policy*. Oxon: Routledge.
- Cable, Daniel, Francesca Gino, and Bradley R. Staats. 2013. Breaking Them in or Eliciting Their Best? Reframing Socialization around Newcomers' Authentic Self-expression. *Administrative Science Quarterly* 58(1): 1-36.

- Chen, Gilad. 2005. Newcomer adaptation in teams: Multilevel antecedents and outcomes. *Academy of Management Journal* 48(1): 101–116.
- Choi, Sungjoo, and Hal G. Rainey. 2010. Managing Diversity in U.S. Federal Agencies: Effects of Diversity and Diversity Management on Employee Perceptions of Organizational Performance. *Public Administration Review* 70(1): 109-21.
- Eisner, Marc Allen, and Kenneth J. Meier. 1990. Presidential control versus bureaucratic power: Explaining the Reagan revolution in antitrust. *American Journal of Political Science* 34(1): 269-287.
- Ellis, Allison M., Talya N. Bauer, Layla R. Mansfield, Berrin Erdogan, Donald M. Truxillo, and Lauren S. Simon. 2015. Navigating uncharted waters: Newcomer socialization through the lens of stress theory. *Journal of Management* 41(1): 203-235.
- Esteban, Joan, Laura Mayoral, and Debraj Ray. 2012. Ethnicity and conflict: An empirical study. *The American Economic Review* 102(4): 1310-1342.
- Feldman, Martha S, and Anne M Khademian. 2000. Managing for Inclusion: Balancing Control and Participation. *International Public Management Journal* 3(2): 149–67.
- Foldy, Erica Gabrielle. 2004. Learning from Diversity: A Theoretical Exploration. *Public Administration Review* 64(5): 529-38.
- Freidson, Eliot. 2001. *Professionalism, the Third Logic: On the Practice of Knowledge*. Chicago: University of Chicago Press.
- Goddard, Yvonne L., Roger D. Goddard, and Megan Tschannen-Moran. 2007. A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers college record* 109(4): 877-896.
- Harris, Douglas. 2011. Value-added measures and the future of educational accountability.

Science 333: 826-827.

Harrison, David A., Kenneth H. Price, Joanne H. Gavin, and Anna T. Florey. 2002. Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. *Academy of Management Journal* 45(5):1029-1045.

Hong, Lu and Scott E. Page. 2001. Problem solving by heterogeneous agents. *Journal of Economic Theory* 97(1): 123-163.

Hong, Lu and Scott E. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Science*. 101(46): 16385-16389.

Hornsey, Matthew J., Tim Grice, Jolanda Jetten, Neil Paulsen, and Victor Callan. "Group-directed criticisms and recommendations for change: Why newcomers arouse more resistance than old-timers." *Personality and Social Psychology Bulletin* 33(7): 1036-1048.

Horowitz, Sujin K. and Irwin B. Horowitz. 2007. The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management* 33(6): 987-1015.

Hvidman, Ulrik and Simon Calmar Andersen. 2014. Impact of Performance Management in Public and Private Organizations *Journal of Public Administration Research and Theory* 24(1): 35-58.

Joardar, Arpita, Tatiana Kostova, and Elizabeth C. Ravlin. 2007. An experimental study of the acceptance of a foreign newcomer into a workgroup. *Journal of International Management* 13(4): 513-537.

Jehn, Karen A., Gregory B. Northcraft, and Margaret A. Neale. 1999. Why Differences Make a

- Difference: A Field Study of Diversity, Conflict and Performance in Workgroups. *Administrative Science Quarterly* 44(4): 741-763.
- Joshi, Aparna And Hyuntak Roh. 2009. The Role of Context in Work Team Diversity Research: A Meta-Analytic Review. *Academy of Management Journal*. 52(3): 987-1015.
- Kane, Aimée A. 2010. Unlocking Knowledge Transfer Potential: Knowledge Demonstrability and Superordinate Social Identity. *Organization Science* 21(3): 643-660.
- Kammeyer-Mueller, John, Connie Wanberg, Alex Rubenstein, and Zhaoli Song. 2013. Support, undermining, and newcomer socialization: Fitting in during the first 90 days. *Academy of Management Journal* 56(4): 1104-1124.
- Levitt, Steven D. and John A. List. 2011. Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments. *American Economic Journal: Applied Economics* 3(1): 224–238.
- Loyd, Denise Lewin, Cynthia S. Wang, Katherine W. Phillips, and Robert B. Lount Jr. 2013. Social category diversity promotes premeeting elaboration: The role of relationship focus. *Organization Science* 24(3): 757-772.
- Mannix, Elizabeth, and Margaret A. Neale. 2005. What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*. 6(2): 31-55.
- Milliken, Frances J., and Luis L. Martins. 1996. Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups. *The Academy of Management Review* 21(2): 402-433.
- Oberfield Zachary. 2014. Accounting for time: Comparing temporal and atemporal analyses of the business case for diversity management . *Public Administration Review* 74(6): 777 –

89.

OECD. 2014. *TALIS 2013 Results: An International Perspective on Teaching and Learning*,

Paris: OECD.

Ospina, Sonia. *Illusions of Opportunity: Employee Expectations and Workplace Inequality*.

Cornell University Press, 1996.

Ospina, Sonia M., and Angel Saz-Carranza. 2010. Paradox and Collaboration in Network

Management. *Administration & Society* 42(4): 404–40.

Parrotta, Pierpaolo, Dario Pozzoli, and Mariola Pytlikova. 2014. The nexus between labor

diversity and firm's innovation. *Journal of Population Economics* 27(2): 303-364.

Pelled, Lisa Hope, Kathleen M. Eisenhardt, and Katherine R. Xin. Exploring the black box: An

analysis of work group diversity, conflict and performance. *Administrative science*

quarterly 44(1): 1-28.

Phillips, Katherine W., Katie A. Liljenquist, and Margaret A. Neale. 2009. Is the pain worth the

gain? The advantages and liabilities of agreeing with socially distinct

newcomers. *Personality and Social Psychology Bulletin* 35(3): 336-350.

Pitts, David. 2011. A little less conversation, a little more action: Using empirical research to

promote social equity. *Journal of Public Administration Research and Theory*, 21(Suppl.

1), i77-i82.

Polzer, Jeffery, Laurie Milton, and William Swarm. 2002. Capitalizing on diversity:

Interpersonal congruence in small work groups. *Administrative Science Quarterly* 47(2)

296-324.

Pounder, Diana G (ed). 1998. *Restructuring Schools for Collaboration: Promises and Pitfalls*.

Albany, NY: SUNY Albany.

- Quick, Kathryn S., and Martha S. Feldman. 2011. Distinguishing Participation and Inclusion. *Journal of Planning Education and Research* 31(3): 272 – 290.
- Rink, Floor and Naomi Ellemers. 2008. Diversity, newcomers and team innovation. In B. Mannix, M. Neale, & K. Phillips (Eds.), *Diversity & groups. Research on managing groups and teams* (pp. 110–135). Stamford, CT: JAI Press.
- Rink, Floor, Aimée A. Kane, Naomi Ellemers, and Gerben Van der Vegt. 2013. Team receptivity to newcomers: Five decades of evidence and future research themes. *The Academy of Management Annals* 7(1): 247-293.
- Sabharwal, Meghna Helisse Levine, Maria D’Agostino. 2016. A Conceptual Content Analysis of 75 Years of Diversity Research in Public Administration. *Review of Public Personnel Administration* <https://doi.org/10.1177/0734371X16671368>
- Schwieren, Christian and Ursula Glunk. 2008. Mechanisms underlying nationality-based discrimination in teams: A quasi-experiment testing predictions from social psychology and microeconomics. *Small Group Research*, 39(6): 643–672.
- Schippers, Michaéla C., Deanne N. Den Hartog, Paul L. Koopman, and Janique A. Wienk. 2003. Diversity and team outcomes: The moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity. *Journal of Organizational Behavior* 24(6) 779-802
- Solinger, Omar N., Woody van Olffen, Robert A. Roe, and Joeri Hofmans. 2013. On becoming (un) committed: A taxonomy and test of newcomer onboarding scenarios. *Organization Science* 24(6): 1640-1661.
- Srikanth, Kannan, Sarah Harvey, and Randall Peterson. 2016. A Dynamic Perspective on Diverse Teams: Moving from the Dual-Process Model to a Dynamic Coordination-based

- Model of Diverse Team Performance. *The Academy of Management Annals* 10(1): 453-493.
- Tajfel, Henri, and John Turner. 1979. An integrative theory of group conflict. In W.G. Austin and S. Worchel (eds.), *The social psychology of intergroup relations*, 33–47. Monterey, CA: Brooks/Cole.
- Van der Vegt, Gerben S., and Onne Janssen. 2003. Joint impact of interdependence and group diversity on innovation. *Journal of Management* 29(5): 729-751.
- Watson, Warren E., Kamallesh Kumar, and Larry K. Michaelsen. 1993. Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. *Academy of Management Journal* 36(3): 590-602.
- Webber, Sheila Simsarian, and Lisa M. Donahue. 2001. Impact of highly and less job-related diversity on work group cohesion and performance: A meta-analysis. *Journal of management* 27(2): 141-162.
- Williams, Katherine Y., and Charles A. O'Reilly. 1998. Demography and diversity in organizations: A review of 40 years of research. *Research in organizational behavior* 20: 77-140.
- Wilson, James Q. 1989. *Bureaucracy: What Government Agencies do and Why they Do it*. New York: Basic Books.
- Zhu, David H., Wei Shen, and Amy J. Hillman. 2014. Recategorization into the in-group: The appointment of demographically different new directors and their subsequent positions on corporate boards. *Administrative Science Quarterly* 59(2): 240-270.

Østergaard, Christian R., Bram Timmermans, Kari Kristinsson. 2011. Does a different view create something new? The effect of employee diversity on innovation. *Research Policy Volume* 40(3): 500–509.

Table 1. The effect of newcomers on team reflection processes

	(1) Discuss the development of specific students' learning with other teachers or other employees at the school	(2) Evaluate my teaching with other teachers or other employees at the school	(3) Observe the teaching of other teachers and give feedback	(4) Problem solving index
Control group	ref.	ref.	ref.	ref.
Educationally similar	0.147 (0.100)	0.385** (0.119)	0.497** (0.124)	0.462** (0.133)
Educationally diverse	0.137 (0.101)	0.200 (0.118)	0.386** (0.119)	0.322* (0.129)
Constant	3.945** (0.0720)	2.899** (0.0752)	1.928** (0.0595)	-0.233* (0.0836)
Observations	499	499	499	499
R-squared	0.007	0.029	0.056	0.041

Standard errors clustered at school level in parentheses. Differences between socially similar and socially diverse treatment groups are not statistically significant at the 5% level.

** p<0.01, * p<0.05

Table 2. The effect of social distinctiveness on newcomer acceptance

	(1) Competencies	(2) Cooperation
Educationally similar	ref.	ref.
Educationally diverse	-0.701** (0.199)	-0.471* (0.199)
Constant	0.325** (0.117)	0.213 (0.120)
Observations	242	241
R-squared	0.122	0.055

School clustered standard errors in parentheses.

** p<0.01, * p<0.05

APPENDIX

1. Description of the randomization procedure

The Danish Ministry of Education invited all 98 Danish municipalities to participate with their schools in a randomized controlled field experiment regarding co-teaching. 68 municipalities applied. Among the applicants, 12 municipalities were selected including 142 schools to compare the effect of educational background on new hires. The municipalities were selected intentionally to create diversity in terms of geography and school size. Within each of these two groups, schools were stratified based on the predicted test score average of the present cohort of 6th grade students. Within each strata, schools were randomly assigned the treatment conditions. Figure A1 is a diagram of the experimental randomization procedure.

Figure A1. The Experimental Randomization Procedure

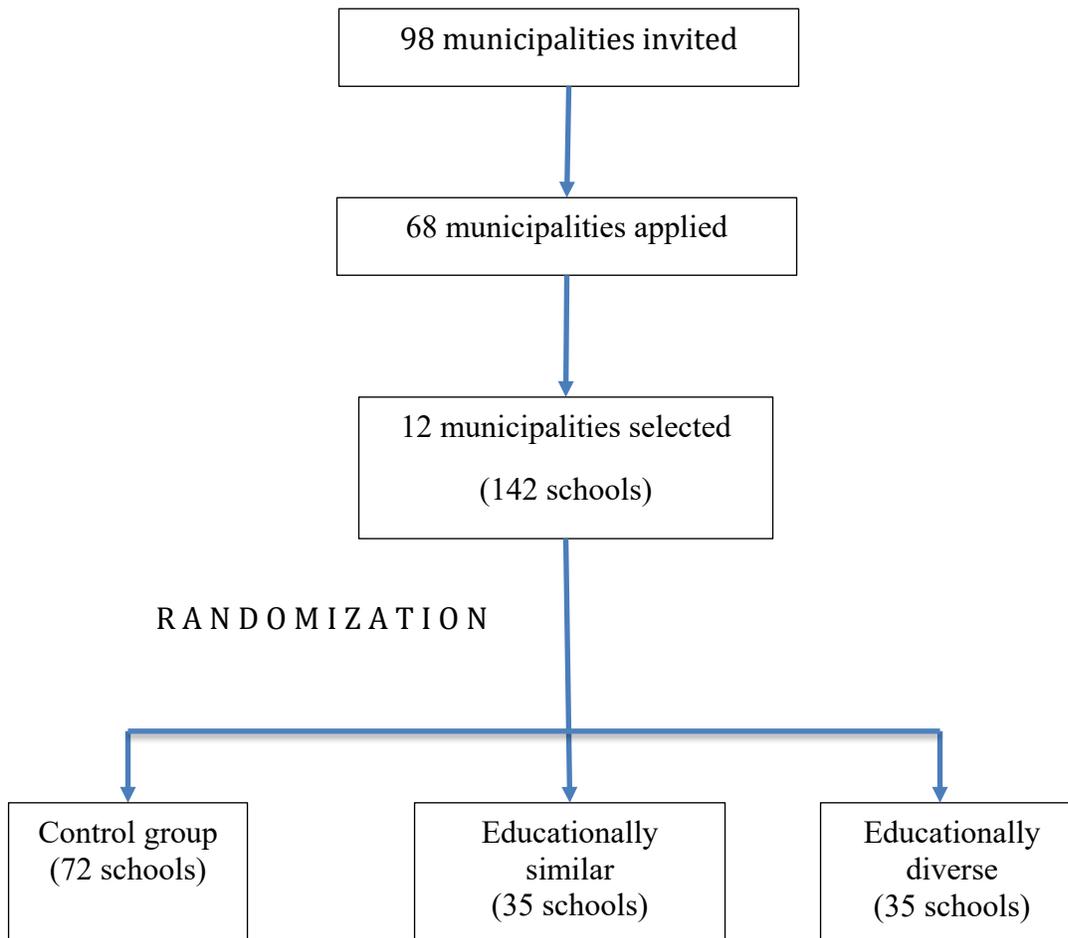


Table A1. Balance of the experimental groups

Table A1 shows that the experimental groups are balanced on a wide range of observed variables.

	Means and percentages			Balancing tests (p-values)		
	(1) Control group	(2) Ed. similar	(3) Ed. diverse	(4) Control = Similar	(5) Control = Diverse	(6) Similar = Diverse
6th grade-cohort: Behavior and achievement						
<i>Pre-tests</i>						
SDQ score 6th grade, autumn 2012	9.068	9.739	9.301	0.098	0.581	0.202
Reading 4th grade, spring 2011	-0.006	0.038	0.051	0.489	0.393	0.840
Math 3rd grade, spring 2010	-0.010	-0.019	0.079	0.889	0.213	0.217
<i>Previous 6th grade cohorts</i>						
Reading 6th grade, spring 2012	-0.084	0.041	-0.032	0.079	0.496	0.310
Reading 6th grade, spring 2011	-0.024	-0.055	-0.025	0.704	0.992	0.698
Reading 6th grade, spring 2010	-0.022	-0.011	-0.008	0.857	0.847	0.963
Math 6th grade, spring 2012	0.026	0.048	-0.012	0.815	0.685	0.537
Math 6th grade, spring 2011	0.007	0.028	0.047	0.816	0.644	0.832
Math 6th grade, spring 2010	-0.026	0.044	-0.032	0.383	0.942	0.370
6th grade-cohort: School size						
No. of students	60.66	60.12	56.39	0.917	0.509	0.536
No. of classes	2.87	2.69	2.65	0.446	0.423	0.844
No. of students per class	21.69	22.42	21.12	0.348	0.435	0.127
6th grade-cohort: Student characteristics						
ADHD and similar diagnoses	0.030	0.030	0.034	0.950	0.600	0.588
Special education needs	0.104	0.123	0.152	0.299	0.024	0.218
Non-Western immigrants	0.019	0.014	0.020	0.395	0.788	0.211
Non-Western descendents	0.107	0.092	0.086	0.650	0.525	0.821
Single mothers	0.239	0.243	0.227	0.836	0.564	0.451
Mothers' age at birth (years)	29.021	29.166	28.920	0.657	0.780	0.489
Mothers' income (10,000 DDK)	24.856	25.293	24.948	0.751	0.948	0.811
Both parents, compulsory schooling only	0.105	0.105	0.084	0.985	0.229	0.272
Both parents, no college degree	0.593	0.592	0.566	0.967	0.480	0.474
Fathers' employed	0.795	0.792	0.796	0.892	0.993	0.898
6th grade-cohort: Teacher characteristics						
Male	0.290	0.413	0.369	0.023	0.217	0.517
Age	41.706	43.755	44.554	0.339	0.135	0.665
Principal survey						
Male	0.650	0.529	0.580	0.345	0.592	0.702
Quality of teaching staff (score 1 to 5)	4.021	3.844	4.007	0.179	0.911	0.187
Quality of teaching staff: Excellent	0.132	0.057	0.105	0.322	0.756	0.500
<i>Priorities: Which of the following three are the most important at your school?</i>						
Students' national test scores	0.078	0.048	0.000	0.622	0.091	0.242
Inclusion of students with special needs in the ordinary classroom teaching	0.204	0.144	0.153	0.542	0.597	0.920
Students' general wellbeing at school	0.718	0.808	0.847	0.412	0.209	0.684

No. of students	1,789	1,814	1,610	3,603	3,399	3,424
-----------------	-------	-------	-------	-------	-------	-------

Note: Columns 1-3 present students average characteristics by experimental group. Columns 4-6 present p-values from pairwise balancing tests of equal average characteristics. Balancing tests are performed by a series of t-tests from regressing each covariate on the treatment indicator with clustering at the school level. School level is determined by the student's initial school allocation (August 2012). Student and parental characteristics are obtained from register data and survey data collected before the beginning of the intervention. Teacher and principal characteristics are from the principal pre-survey. Test scores are standardized in the population with mean zero and a standard deviation of one within each year, grade and subject. SDQ score ranges from 0 to 40.