ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Worren, Nicolay; Christiansen, Tore; Soldal, Kim Verner

Article

Using an algorithmic approach for grouping roles and sub-units

Journal of Organization Design

Provided in Cooperation with:
Organizational Design Community (ODC), Aarhus

Suggested Citation: Worren, Nicolay; Christiansen, Tore; Soldal, Kim Verner (2020) : Using an algorithmic approach for grouping roles and sub-units, Journal of Organization Design, ISSN 2245-408X, Springer, Cham, Vol. 9, Iss. 1, pp. 1-19, https://doi.org/10.1186/s41469-020-0069-2

This Version is available at:
https://hdl.handle.net/10419/252156

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by/4.0/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Worren et al. Journal of Organization Design (2020) 9:8
https://doi.org/10.1186/s41469-020-0069-2

Journal of Organization Design
a SpringerOpen Journal

RESEARCH Open Access

Using an algorithmic approach for grouping roles and sub-units

Nicolay Worren[1*], Tore Christiansen[2] and Kim Verner Soldal[2]

* Correspondence: nicolay.worren@nmbu.no
[1]School of Economics and Business, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Ås, Norway
Full list of author information is available at the end of the article

Abstract

A key task in organization design is to group elements (e.g., roles) into sub-units (e.g., teams or departments). This task is computationally challenging as one must take into consideration a potentially large number of interdependencies between the elements. It also requires data about work processes in the organization, which are rarely present. We have initiated a research program that aims at developing a tool—Reconfig—to improve grouping decisions. It first collects data from employees about their working relationships (i.e., interdependencies) and then uses a computer algorithm to cluster the data in the most optimal manner. The clustered solution represents the formal structure that minimizes coordination costs by grouping the most highly interdependent elements together. We describe the tool, report on two pilot applications, and discuss both the future potential and limitations of the approach.

Keywords: Organization design, Clustering algorithm, Design structure matrix (DSM), Interdependencies, Software tool

Introduction

The early contributors to our field identified the topic of grouping as a key element in organization design:

> Given a general purpose for an organization, we can identify the unit tasks necessary to achieve that purpose (…) The problem is to group these tasks into individual jobs, to group the jobs into administrative units, to group the units into larger units, and finally to establish the top level departments – and to make these groupings in such a way as to minimize the total cost of carrying out all of the activities (March & Simon, 1958, p. 41)

We believe, however, that this is a challenging task to carry out in practice. It would have been simple if it was merely a question of categorization—of placing one set of roles (say, those related to software in a technology firm) in one sub-unit, and another set (say, those related to hardware) in another sub-unit.[1] A key characteristic of most roles (and sub-units) in organizations is that they are *interdependent*, first and foremost in relation to the work processes that they carry

[1]An earlier version of this paper was presented at the Journal of Organization Design conference in Sandbjerg, Denmark, 29th-31st of May, 2018.

Springer Open

© The Author(s). 2020 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

out, but also in relation to other types of interdependencies, such as resources and governance relationships (Worren 2018). Handling these interdependencies requires interchange of information and coordination, leading to coordination costs (Galbraith 1974; Nuñez et al. 2009). Having a large number of interdependencies across sub-units may also make it difficult to define clear accountability for a complete deliverable (Kilmann 1983). Hence, in the organization design process, it is not sufficient to group the organizational elements (roles, sub-units, etc.) based on a quick categorization; one must take into consideration how one should group the elements in order to minimize coordination costs.

Grouping occurs at multiple levels. A project manager responsible for a large engineering project will need to decide who should work in which team within the project (i.e., he/she will find a way to allocate roles to teams). Similarly, an executive vice president may reconsider how the business units within the business area are defined, and conclude that some units should be split and that others should be combined. A CEO may similarly evaluate the overall organization of the firm and change the number of business areas or alter their composition (i.e., business units may be moved between the business areas and regrouped).

A failure to perform this grouping task appropriately leads to a lack of alignment between the formal structure and the work processes in the organization. Studies have shown that interdependencies are sometimes underestimated (Sherman and Keller 2011) leading to a lack of integration and coordination between roles and activities. In other cases, they are overestimated, leading to a lack of separation and to excessive teamwork or unnecessary collaboration (Cross and Gray 2013).

We believe that there are two challenges when performing this grouping task in practice. The first is a lack of access to relevant data. The functional labels such as the titles of positions ("Financial controller") or sub-units ("Accounting and Finance") are readily available and documented in formal organization charts, whereas interdependencies are for the most part implicit. This was recently confirmed in a survey with 176 organization design consultants reported in this journal (Worren et al. 2019). The survey sought to identify the key challenges

Fig. 1 Item result from survey among organization design practitioners ($N = 176$) (Worren et al. 2020)

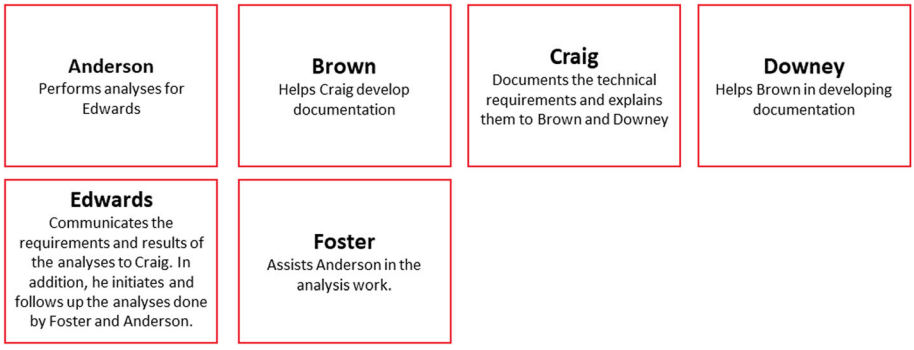

Fig. 2 Role descriptions used in study of role grouping (adapted from Worren et al. 2020)

that consultants encounter when assisting clients in re-designing their organizations. The respondents were asked to evaluate 25 different elements of the re-design project, from scoping the engagement with the client to planning and implementing the new organizational model. One of the findings was that a large proportion (60%) indicated that it is frequently or always a challenge in to understand how people in the client organization collaborate or exchange information across units (Fig. 1).

A second challenge is the process of grouping elements itself. Even when one has access data about how people interact, it is difficult to group different elements while keeping in mind the interdependencies. In an organization design project, there are typically hundreds of interdependent elements that need to be sorted or clustered into the best possible configuration. We believe that this task is cognitively demanding and subject to various psychological limitations and biases. To study this in more detail, Worren et al. (2020) conducted a study where they asked participants to group interdependent roles into small teams. They simplified the task by reducing the number of elements to as little as six roles and by creating problems with a clear-cut optimal solution (they also used tasks with nine and twelve roles). They started by showing participants a set of role descriptions, such as those shown in Fig. 2. The task is to find out how to group the roles into the teams. The participants were informed that there was a rule in this organization that there should be a maximum of three people in each team. The results show that even with as few as six roles and with no time limit, 25% of the participants did not succeed in identifying the optimal

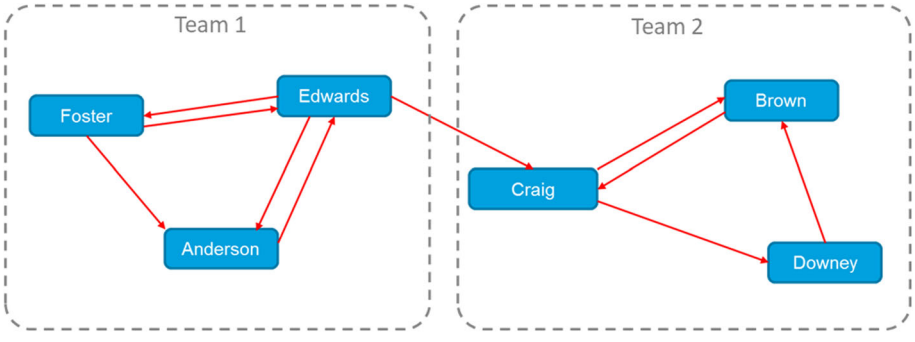

Fig. 3 Optimal solution for the task shown in Fig. 2 (adapted from Worren et al. 2020)

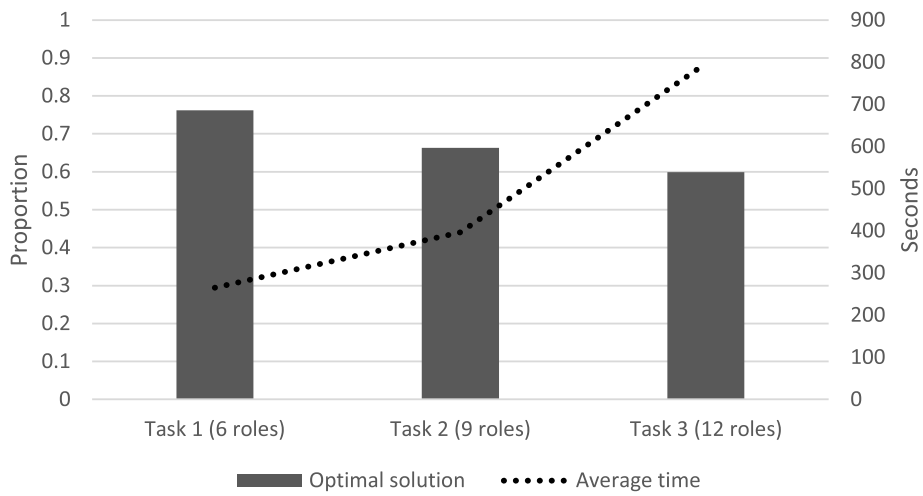

Fig. 4 Proportion of participants who identify the optimal solution and time usage across the tasks with different levels of complexity (i.e., number of elements) (Adapted from Worren et al. 2020)

solution. The proportion of participants that identified the optimal solution decreased when the number of elements was increased (Fig. 3). Worren et al. (2020) also found that participants used significantly more time to solve the task as the number of roles increased (see Fig. 4). We would expect that even fewer participants would identify the optimal solution in a more realistic task with dozens or hundreds of elements. We have used variations of this task in courses with both students and managers. The results have been similar. During the debriefing, we typically find that participants agree that the optimal solution is the best one, even if they did not identify it themselves.

The reason why clustering tasks are difficult is related to the number of potential permutations of sets of elements, which increases rapidly with size. Three roles may be allocated to two teams in three different ways:

Team 1	Team 2
A B	C
A C	B
B C	A

If instead there are six roles that need to be allocated into two teams, this may be done in 41 different ways.[2] In general, the number of ways to select a sub-unit of k members from a unit of n members is given by the binomial coefficient:

$$\frac{nP\mathrm{k}}{k!} = \frac{n!}{(n-k)!\cdot k!} = \binom{n}{k}$$

The result of introducing additional elements on the number of permutations is shown in Fig. 5.

[2]Since the number of members in the subunit is arbitrary we must sum for all k to get the numbers in the graph. But since the remaining (n-k) members can be organized into more than one subunit, the total number of possible combinations is even larger than given here.

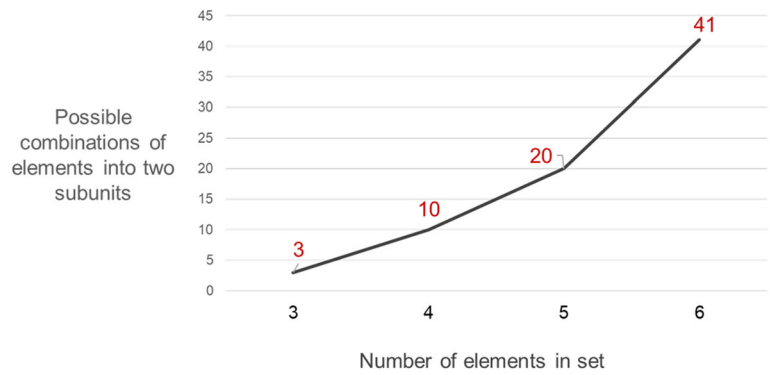

Fig. 5 The number of possible combinations of elements into two sub-units as a function of the number of elements in a set

Given the existence of heterogenic interdependencies between the elements, the various permutations (clusterings) will produce solutions with varying degrees of effectiveness. Identifying the optimal clustering would thus require the decision maker to search the entire space of permutations and calculate the degree of optimality of each. Since this task exceeds the cognitive capacity of humans, a more likely strategy is that decision makers retort to using a heuristic when developing an organizational design. For example, they may select a grouping based on another criterion, such employees' functional specialization, rather than taking into account interdependencies related to work processes (Worren et al. 2020).

The success rate of organization design processes is generally quite low. Scholars who study strategic decisions (including those pertaining to organization re-design) have estimated the success rate to be around 50% (e.g., Nutt 1999). The two factors discussed above—limited access to relevant data and difficulty in processing the information (if present)—may partly explain why it is difficult to make effective decisions. Important information is not collected or is ignored, or elements are combined in a suboptimal manner. The end result is that ineffective organizational models are being introduced, for example, organizational models that do not reflect the actual work processes in the organization, leading to increased coordination cost, conflicting roles, and unclear unit mandates. A recent example is provided in a case study of a reorganization of a large European food production company (Livijn 2019). The top executives who introduced the new organizational model presented it as a functional organization (i.e., a structure with unitary reporting relationships) with clear roles and responsibilities. The middle managers who were implementing the new model discovered that it was in fact a full blown matrix (i.e., with multiple reporting relationships), with unclear interfaces and missing co-ordination mechanisms. Managers and employees in large firms generally perceive their organization to be highly complex; in one survey, only 1% of respondents indicated that their organization was "not complex at all" (Economist Intelligence Unit 2015). The same survey also asked the respondents to identify the effects of organizational complexity. The most frequently mentioned effects were slower decision making, an excessive amount of time spent on coordination, and lower employee morale (ibid.).

Table 1 High-level functional requirements for organization design tool

#	Functional requirement	Explanation
1	Support collection of data from the organization	The tool should make it possible to collect data about work process-related interdependencies between roles
2	Allow visualization of interdependencies	The tool should include visual maps and provide the decision maker with an overview of the role interdependencies
3	Generate grouping of elements (roles, sub-units etc.)	The tool should provide a suggested grouping (or clustering) of the data that minimizes coordination cost, and thus indicate an improved way of organizing roles/sub-units

We believe that an effective organization design tool would need to fulfill three main functional requirements, related to data collection, visualization, and analysis (see Table 1). In the following, we describe in more detail how we are developing a tool—Reconfig—that fulfills these requirements and we discuss our experience from using a prototype version of the tool.

An algorithmic approach to grouping

There are numerous algorithms that will take a set of interdependent elements and produce a clustering (Xu and Tian 2015). We use a genetic (or evolutionary) algorithm (Holland 1992) developed by Soldal (2012). The input to the algorithm is a square matrix with elements and their interdependencies called Design Structure Matrix (DSM) (Browning 2001). The output is a list of numbers that represents the groups (i.e., clusters) that each element in the DSM (i.e., role or sub-unit) has been allocated to.

The genetic algorithm goes through a series of steps as outlined in Fig. 6. It first generates a set of random solutions whose performance (fitness) are then evaluated and ranked according to a fitness function. The 50 highest scoring solutions are then selected as parents and generate 200 offspring that are composed of combinations of their genetic code. Random genes in the new offspring are then mutated by randomly changing values. The offspring is finally subjected to another round of fitness evaluation, and the 50 offspring with the highest fitness become parents for the next generation, and so on. The process terminates either after a set number of cycles/generations or when the solution is stable (for a more detailed explanation of genetic algorithms, see Goldberg 2009).

The fitness function that we use was developed by Yu et al. 2007) and seeks to identify an improved grouping of the data (i.e., one that maximizes within-cluster interdependencies and minimizes between-cluster interdependencies):

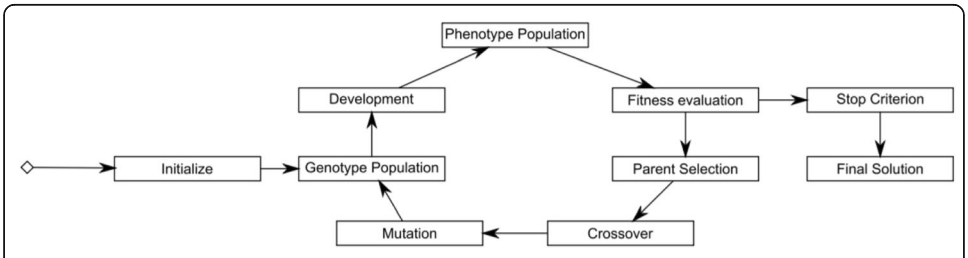

Fig. 6 The flow of the genetic algorithm (adapted from Soldal 2012)

$$f_{DSM}(M) = (1-\alpha-\beta)\left(n_c \log n_n + \log n_n \sum_{i=1}^{n_c} cl_i\right) + \alpha\left[|S_1|\,(2\log n_n + 1)\right] + \beta\left[|S_2|(2\log n_n + 1)\right]$$

Where:

n_c= Number of clusters in the DSM

n_n= Number of rows/columns in the DSM (i.e., elements such as number of roles)

cl_i= Number of nodes in cluster i

S_1= Sum (Type 1 error) = "Errors of omission" (failing to include an interdependent element in a cluster)

S_2= Sum (Type 2 error) = "Errors of commission" (including a non-interdependent element in a cluster)

α and β= Weights between 0 and 1

The objective in the above equation is to find a solution that minimizes f_{DSM} (i.e., coordination cost). For a given model (M), the result is the solution with the lowest coordination cost for the chosen parameters.[3]

Rationale for selecting the genetic algorithm

There are multiple algorithms in the field of network analysis that are intended to find the most optimal subgroups of elements, also called graph partitioning, such as the Kernighan-Lin algorithm, Fiduccia-Mattheyse's algorithm, and spectral clustering (for useful reviews, see Buluç et al. 2016; Schaeffer 2007) These algorithms generally use heuristics to reduce the computation cost of scaling the problem for large problems. However, traditional graph partitioning methods are intended to solve a predefined mathematical problem with limited consideration made for the applied use of the method. In contrast, evolutionary methods allow for a very high degree of flexibility. The fitness function that evaluates each generated solution can include any metrics that can be evaluated mathematically, and multiple metrics can be combined into one fitness function. This means that one is able to specify a varied number of evaluation criteria that result in graph partitions with particular characteristics that matter to organizations. For example, one can specify a maximum number of clusters and run the analysis based on this constraint. One can also adjust the alpha and beta weights in the fitness function (see the equation above) and thereby prioritize whether one minimizes interdependencies *between* clusters (alpha) or minimizes the number of non-interdependent elements *within* a cluster (beta). Finally, one can also select interdependencies related to a sub-set of activities or business processes and run the algorithm on these only (e.g., to consider which organizational model that would be the most optimal when including only sales & marketing processes).

The cost of this flexibility is that genetic algorithms are less scalable than traditional graph partitioning algorithms. However, for organization re-design projects, even in

[3]The alpha and beta weights are set at 0.3 each as a default, but may be adjusted. By increasing the alpha weight (and reducing the beta weight correspondingly), one prioritizes reducing the number of interdependencies *between* clusters; by increasing the beta weight (and reducing the alpha weight), one prioritizes reducing the number of interdependencies *within* clusters. In our further work, we aim at testing more systematically the effect of varying the weights, and identifying when it may be appropriate to depart from the default set-up.

large global organizations, the number of roles or units being considered is rarely large enough for scale to be a critical issue (if one reorganizes a large firm, one can limit the respondents to a few representatives from each unit, instead of including all employees). Nor is it always necessary to find the optimal solution. Our focus is to generate graph partitions that meet criteria better than intuition would. Genetic algorithms also provide the flexibility to limit how long an optimization takes. While there may be smaller improvements to gain after running the algorithm for 10,000 generations, it is likely that a usable solution is identified after 4000 generations. By accepting this goal of a "good enough" solution, it will only take hours to calculate the solution for an organization of 500 people (or units), instead of days.

Comparison to social network analysis

Our approach is similar, but not identical, to recent developments in social network analysis. Whereas social network mapping has its roots in sociology (Freeman 2004), our approach is based on methods developed in the engineering sciences (Eppinger and Browning 2012; Sharman and Yassine 2004). As the name suggests, social network analysis originally focused on social relationships, and typically included "affective" variables (e.g., mutual trust and liking among a group of people) (Freeman 2004). More recently, there are also examples of social network studies that focus on more task related or "instrumental" variables such as knowledge transfer (e.g., Hansen 1999). Social network scholars have also started to apply algorithms to detect "community structures." A community structure is similar to a cluster as defined here—"network nodes [that] are joined together in tightly knit groups, between which there are only looser connections" (Girvan and Newman 2002).[4] Nonetheless, there are also some differences. Both the purpose and the data that are included in the analyses differ: The purpose of Reconfig is not to analyze social networks among individuals per se, but to align the formal structure of organizations (i.e., a set of roles) with the work processes. To achieve this, we not only collect and analyze data about the task related interdependencies, but also about the formal structure (i.e., the allocation of roles to different divisions, departments, or teams).

Data collection

The data that are processed by Reconfig are collected from managers and employees of the organization. We currently use a commercially available electronic survey tool (SurveyGizmo), which has been configured specifically for the purpose. It consists of three pages. On the first page, respondents are asked to indicate which activities or work processes they participate in. The list only contains the main activities or work processes in the organization and is thus quite brief (3–10 items). (Yet it is quite important, as it makes it possible to filter the data by activity both when visualizing and analyzing the results.) On the second page, respondents are asked to pick up to seven people that they collaborate with (all employees in the organization are listed and can

[4]Girvan and Newman (2002) propose a measure for the strength of the community structure, as an objective metric for choosing the number of communities into which a network should be divided. This metric, or some variation of it, may be relevant to rank and compare the optimized organization structure resulting from the genetic algorithm in Reconfig.

be found by typing the first letters of their first name). On the third page, the respondents are asked to characterize the relationship with each of the (up to) seven people. They are first asked to indicate the direction, that is, whether they receive or provide something to the other person, or whether they consider it to be a mutually interdependent relationship. They are then asked which activity or work process that it relates to. Finally, they are asked to rate (one a three-point scale) how important the collaboration is for their ability to reach their goals (Fig. 7). The results can be visualized in a DSM as shown in Fig. 8.

Example 1: university

We tested the tool in connection with the reorganization of a university in Norway in 2016. The university had implemented a reorganization the year before, and we were curious about whether the new organization was aligned with the most important interdependencies. An external consulting firm had been hired to help the university president identify a more optimal formal structure. A key concern with the previous structure was that there were too many hierarchical levels, taking into account the university's limited size (Around 5200 students and 1700 staff). The university had 13 main departments, three schools or faculties, and two research centers. Hence, removing the "school level" would result in 15 units reporting to the university president (plus staff functions such as IT, HR, and the library), which was deemed as excessive. The university president indicated that it should be possible to develop a model with only six or seven main units. In other words, a re-grouping was necessary, and a key task was to find an appropriate grouping of the 13 departments and two research centers.

The structure prior to the reorganization in 2016 is shown in Fig. 9. We observe that despite the three large schools, several interdependencies were not contained within the schools. We also note that school 3 consisted of three weakly interconnected departments.

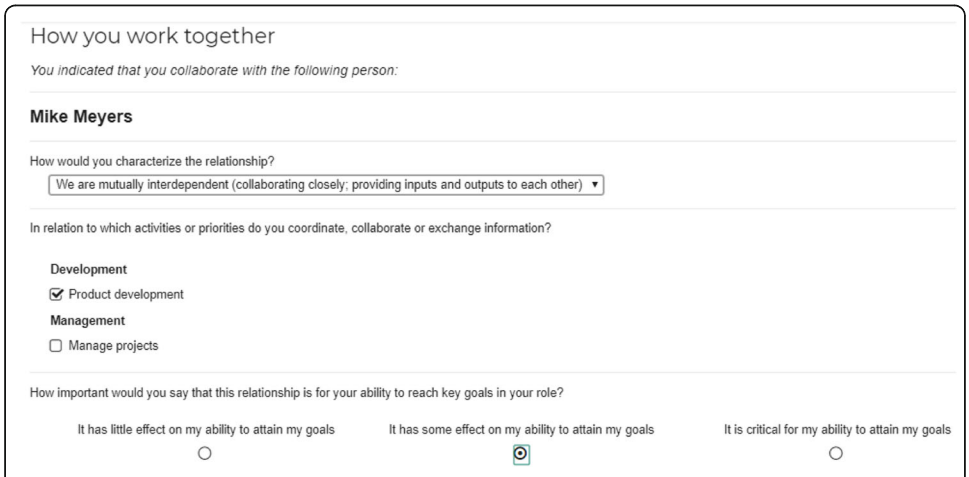

Fig. 7 Screen image from the survey questionnaire used to gather data

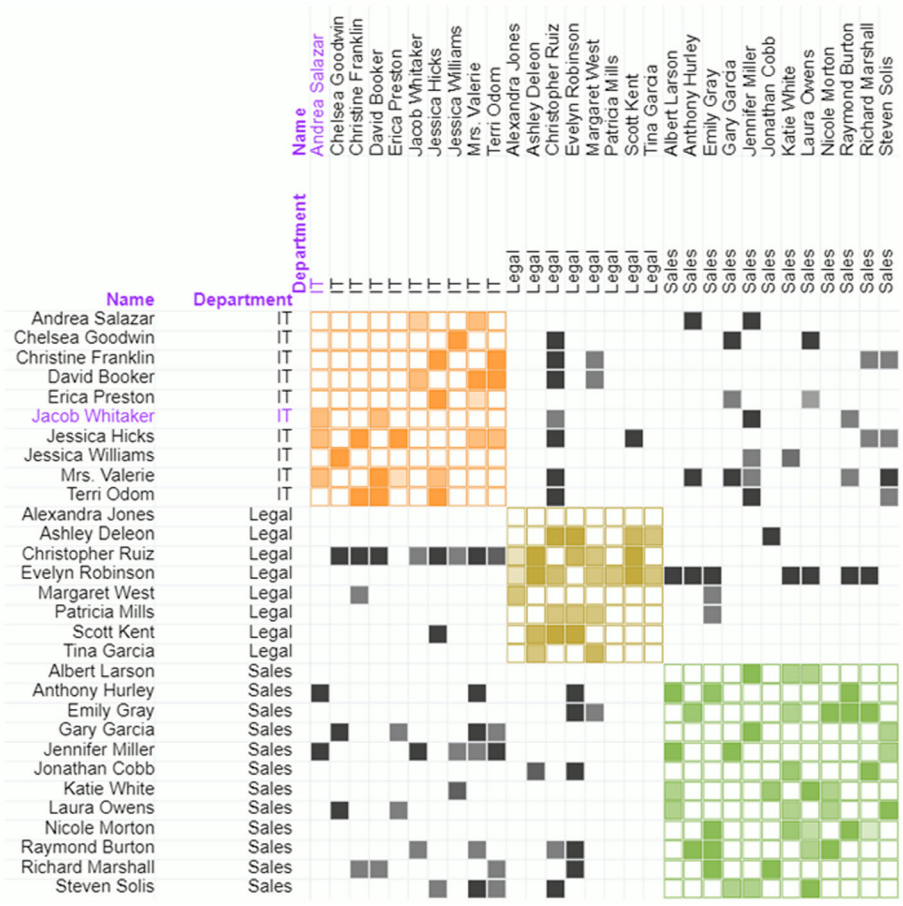

Fig. 8 Screen image of the tool, showing data for the current organization in a DSM format (the names are randomly generated). The color shade indicates the degree of criticality of the interdependency as reported by the respondents.

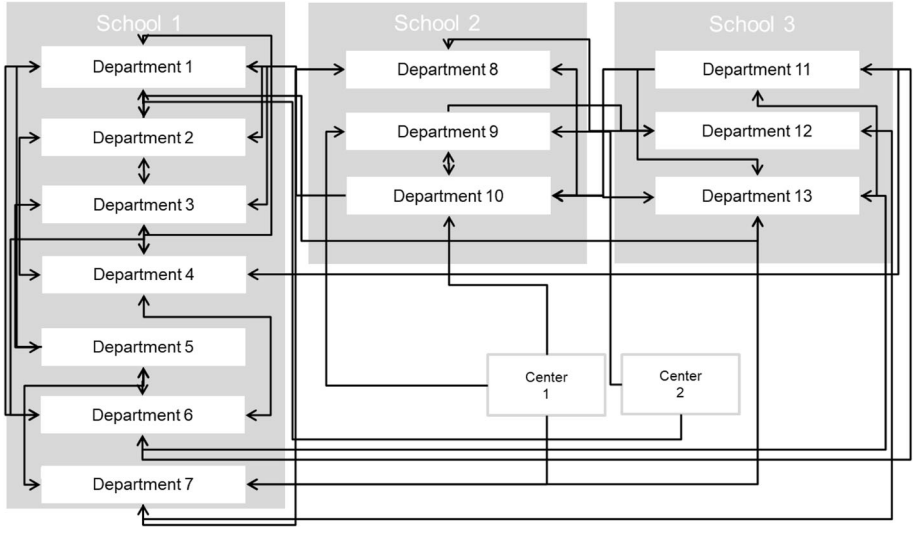

Fig. 9 Organization of university prior to reorganization. The illustration was drawn in PowerPoint based on data from Reconfig

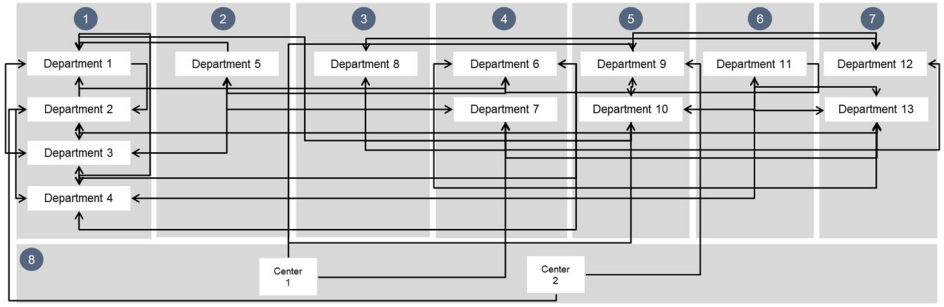

Fig. 10 Organizational structure that was actually implemented in university. The illustration was drawn in PowerPoint based on data from Reconfig

Various grouping criteria had been proposed, but the criterion that was deemed the most important was related to coordination and collaboration with regards to study programs. Whereas collaboration in research projects might occur more or less independently of formal boundaries, it was concluded that the sub-units involved in joint study programs should be formally grouped together. A new organization was introduced during Fall 2016 (Fig. 10). We observe that the grouping does not contain several of the interdependencies. We also see that two units that have no interaction with each other (Departments 12 and 13) are grouped together.

We first mapped the interdependencies (related to the structure prior to the reorganization). This was done by distributing a survey questionnaire to all department heads and asking them to indicate the interfaces they had toward other units. They were also asked to rate the importance of each interface and indicate whether it was related to study programs, research, administration, or other priorities.

Using the fitness function described above (Yu et al. 2007), we then calculated the coordination cost of the organizational structure prior to the reorganization, and compared it to the solution generated by the tool (Fig. 11). Compared to the prior organizational structure, we found that the solution generated by the tool would reduce the coordination cost with 32% (from 149 to 100). Compared with the prior organization, the organizational structure that was actually implemented

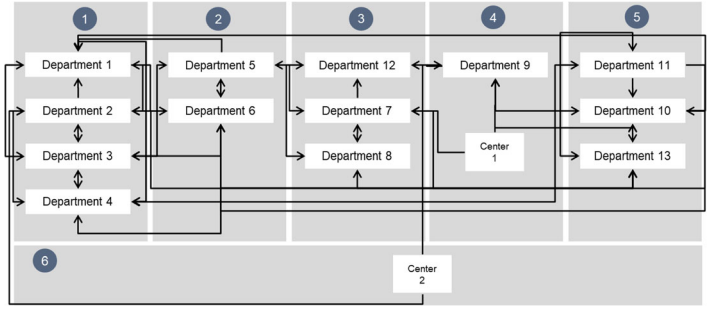

Fig. 11 Suggested tool solution for university. The illustration was drawn in PowerPoint based on data from Reconfig

increased the coordination cost by 23% (from 149 to 184). As can be seen when comparing Figs. 10 and 11, there is only one cluster (i.e., school/faculty) that is similar in the two solutions (cluster 1). The tool solution reduces the overall number of interdependencies outside the clusters, and also avoids grouping unrelated departments.

Example 2: manufacturing firm

Recently, two master level students (Karlsen and Gronvold 2019), supervised by the main author, had the chance to test the tool in a small manufacturing firm with 43 people. The firm is in a growth phase and is considering how to adapt the organizational structure as it hires more employees. It started out with a simple structure (i.e., little or no differentiation between roles and sub-units) (Mintzberg 1983), but had introduced a traditional functional structure with departments for procurement, sales, assembly, soldering, and inventory and shipping. Due to the small size of the organization and the desire to identify the most appropriate team structure, all employees were included in the survey phase. All of the employees (with the exception of three part time employees) filled in the survey questionnaire. The questionnaire asked them to indicate which work activities they participated in and who they collaborated with. As with the study described above in the university, the employees were also asked to indicate the degree of criticality of each relationship with regards to goal attainment.

A Design Structure Matrix (DSM) based on the current organization is shown in Fig. 12. Reading along a row, the marks represent elements (in this case roles) that the

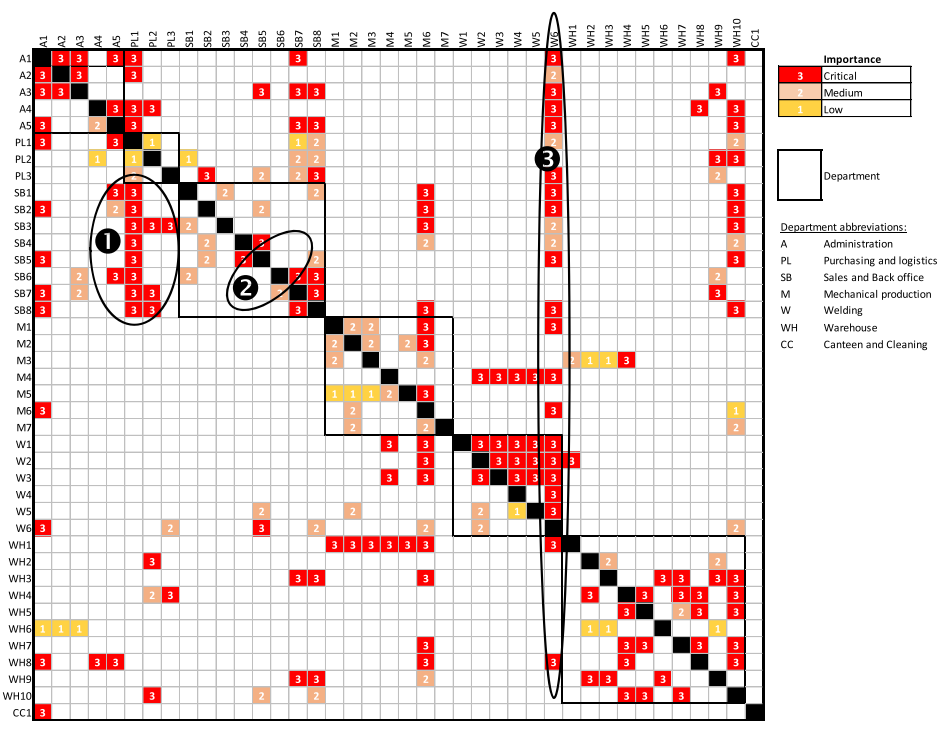

Fig. 12 Design structure matrix (DSM) for the current organization of small manufacturing firm

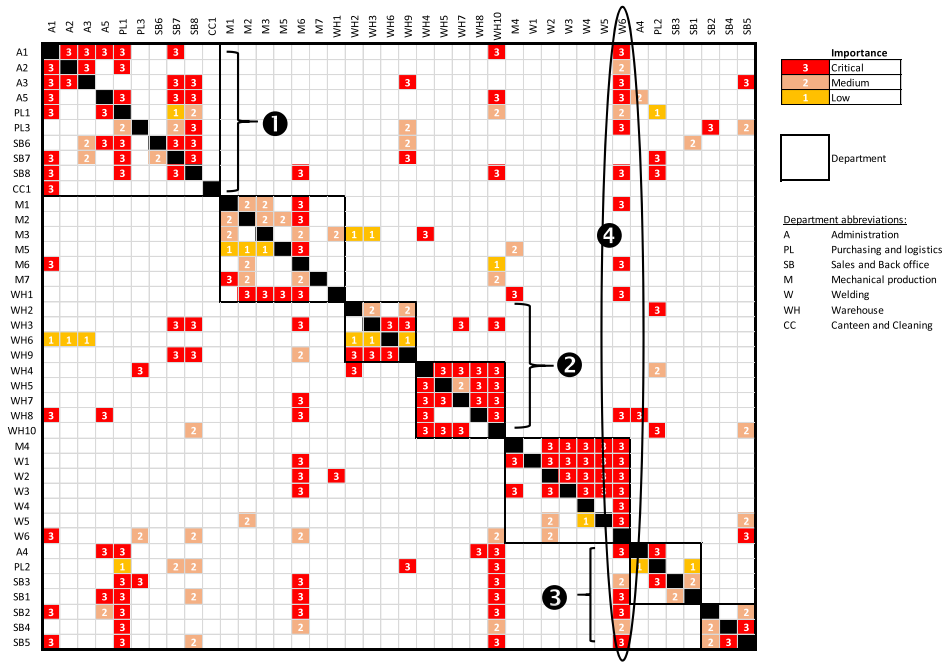

Fig. 13 Design structure matrix (DSM) showing Reconfig's clustered solution for the organization of small manufacturing firm

role on the left of the row needs inputs from (e.g., information, resources). Reading along the column, the marks indicate elements (roles) that the role above the column provides outputs to. The borders inside the matrix indicate which element belongs to which department. The DSM indicates that there is relatively poor alignment between the formal structure and the work processes. First, we see that a number of interdependencies, including ones considered to be of critical importance, are outside the clusters (departments): Of the 229 dependencies mapped, 136 are between roles belonging to different departments (see, for example, the area marked with (1) on Fig. 12). Second, we note that there are a few roles that seem to be weakly connected to each other, although they are placed in the same department (indicated by the white cells inside the clusters) (see the area marked with (2) in Fig. 2). We also note the existence of a few roles that have a large number of interdependencies. In particular, 23 employees (59% of the staff) indicated that they were dependent upon employee W6 in the Welding department (see the area marked with (3) in Fig. 12). Of the 23 employees, 18 are placed in another department than W6.

The five managers of the firm, including the managing director, participated in a session were the results were presented and discussed. They expressed that the analysis was consistent with their understanding of the way people actually performed their jobs today.

The clustered solution is shown in Fig. 13. It reduces the coordination cost (fitness score) from 1113 to 791 (29%). The solution contains seven clusters (i.e., departments/teams). There are three notable differences from the current organizations: (1) A cross-functional department of ten people, representing roles from the administration, purchasing and logistics, and sales and back office; (2) two welding departments (or sub-groups), each tightly connected internally; and (3) two sales and back-office teams.

The clustered DSM clearly shows stronger alignment between process and structure. The number of interdependencies between clusters is reduced by 30 (28%) while the number of interdependencies within clusters is increased with 30 (32%). At the same time, one would probably want to make some adjustments before this solution is translated into a new organization chart. For example, the two sales and back office teams (marked (3)) may be viewed as too small to be managed by separate managers, and may be merged into one department. Furthermore, the re-clustering did not result in a change with regard to W6, which is still connected to 23 other employees (marked (4)). One will need additional information to determine whether this is an appropriate solution or not. On the one hand, it may be the case that this particular employee is overloaded with requests from others, and that either the work processes should be changed (to reduce the demand), or that capacity should be added by allocating one more person to the task carried out by this person. On the other hand, this may also be a coordinating role (a "bus" in DSM terminology) that serves to integrate the work of many others in the organization, in which case it may be preserved as it is.

Discussion

Scholars in the field of organizational theory have largely focused on *describing* organizations, instead of changing or improving them. Reconfig is an example of solution oriented social science (Watts 2017). We are developing a tool for practitioners based on key principles in organizational theory. But the link is not one-way (from theory to practice), but two-way (from theory to practice, and back to theory): In the future, we expect that the development process itself—and the data that we collect from organizations—can be the basis for empirical analysis and further theoretical development.

Our starting point was to conceptualize organization design as a grouping task. It requires access to data about interdependencies between elements (roles, sub-units, etc.) and the ability to process these data to identify the clustering (or grouping of elements) that is the most optimal one. This task is often hampered by a lack of relevant data. Reconfig addresses this need by providing a means for visualizing and analyzing data about internal working relationships. Reconfig thus confers "situational visibility" (Danilovic and Sandkull 2005; Steward 2007). It ensures that decisions can be based on an understanding of how the organization actually functions. Distribution of the survey has the added benefit of involving employees in the organization design process and signaling the intention of management to use a "facts based," rather than intuitive or political, approach to designing the organization.

The clustering algorithm may be viewed as a way of operationalizing the concept of a "modular organization." It identifies a structure consisting of semi-independent units or modules with a maximum number of within-cluster interdependencies and minimal number of between-cluster interdependencies. As explained in Sanchez (1995) and Clark and Baldwin (2000), increasing the degree of modularity is a key strategy for simplifying complex organizations. The key benefits are a reduction in the time and cost of coordination, assuming that coordination is less costly within than between sub-units. But there are also other potential benefits. In a study of Microsoft programming teams, it was found that having more independent and specialized teams increased the quality of the end product. Indeed, it was found that the organizational structure predicted

defects with very high precision (86%) and that it was a better predictor of defects than traditional process and software related metrics (Nagappan et al. 2008). In a simulation study, Fang et al. (2010) found that having semi-isolated subgroups fostered innovation, as long as there were some connections remaining between the groups. The relative isolation of groups shielded the development of divergent ideas within each group, while the connections between groups helped diffuse the most promising ideas across groups.

In this article, we have focused mainly on the analytical basis for the tool and its technical design. However, the intention is not to fully "automate" organization design. On the contrary, we view the utilization of the tool as we would view any intervention in a social system. Even with the best of algorithms, there are several decisions that will need to be made by human decisions makers in such processes, including when to start a review of the current organization and how to define the scope of the project. The data that the tool provide will be presented to and interpreted by managers, who in some cases will choose to disregard some of the results, while relying on others. This may be perfectly legitimate, as the algorithm does not take into consideration all relevant factors. It is also clear that the use of the tool requires a certain level of trust between employees and management (and the advisor/consultant who uses the tool). Employees are asked to participate and provide information about what they do and who they work with. Based on our project so far, we believe it is possible to achieve close to a 100% response rate among employees in the data collection phase. But this requires that managers are able to communicate with employees and explain the purpose of the project. In our experience, employees will participate if they know that their input will be taken into consideration, if they receive some feedback on what the overall results of the survey were, and if they are confident that their personal data will not be misused.

Limitations

Although Reconfig represents a new step toward a more analytical and data-based methodology for organization design, several limitations should also be acknowledged. Some are due to the overall approach being used, while others are due to resource constraints and simplifying assumptions having been made when developing the tool.

First, the tool is intended to support decisions about "horizontal" grouping. Currently, the approach does not incorporate an analysis of the vertical complexity of a proposed design. For example, in the first example described above, the difference in fitness score between the original structure at the university and the new model that was adopted does not reflect the fact that the new model represented a "de-layering" (i.e., one management level was removed). A more complete analysis would need to include data about the vertical structure and take this into consideration when clustering the elements.

Secondly, in collecting data, we essentially collapse two dimensions of interdependency: Interdependencies between individuals (as evidenced by communication flows) and interdependencies that are due to tasks (that are assigned to roles). As is common in organization design theory (Jaques 1989), we distinguish between individuals and the roles they hold. In principle, we do not want to identify individual interdependencies per se. The purpose is not to find

the best mix of individuals (e.g., based on their personality), but the appropriate grouping of roles. The individuals that currently hold the roles may be re-allocated to other roles as a result of the re-design process that may follow. Researchers like Sosa et al. (2004) who study engineering firms have conducted two separate studies of the current organization, one that documents the communication flow, and one that documents that technical (i.e., task) interdependencies. This approach makes it possible to evaluate the degree of correspondence between actual and needed coordination. However, we believe that it is too cumbersome in practice to collect data at two levels. Hence, we only collect information from individuals about their communication flow/working relationships. On the other hand, we do focus on task-related interdependencies and do not simply ask respondents who they prefer to work with; the questionnaire first asks respondents to list activities or work processes they contribute to, and then to list who they depend on or deliver to in each of the activities or processes that they have listed.

Thirdly, the tool does not map all aspects of interdependency. At least three aspects of interdependency are discussed in the literature: direction, criticality, and uncertainty (or predictability). We do cover the direction, as shown above, and also collect data about criticality.[5] However, we do not collect data about the uncertainty of the interdependency. This is a potentially important limitation, as a key assumption in the literature is that a less uncertain interdependency (despite being of high importance) can more easily be formalized and handled by means of planning and documentation (Van de Ven et al. 2006). This suggests that interdependencies of unequal uncertainty should not be treated equally by the clustering algorithm (it should prioritize the interdependencies of high uncertainty).

What Reconfig does is to optimize the clustering or grouping of elements. Yet we do not know precisely the cost of a *non-optimal* grouping, or if one considers optimality as a continuum, at what point a non-optimal grouping starts having a negative effect on performance. Are there critical limits on the amount of coordination an organization can handle before suffering performance degradation? This challenge is shared with most other theories in the field of organization and management: There are few or no norms of what constitutes a "normal" level against which one can assess the current situation. As pointed out by Jaques (2002), this would be similar to practicing medicine without the concepts of normal temperature, blood pressure, pulse rate, and so on.

Reconfig relies on a "bottom up" process where data about work processes are used to determine the appropriate grouping of elements. In practice, we believe that such a "bottom up" method should not be used in isolation but combined with a "top down" approach deriving the appropriate design from an analysis of the organization's mission, functions/capabilities and strategic goals (see Worren 2018). As an example, a "top down" analysis may sometimes identify sub-units that need to be separated (e.g., to due conflicting goals or mandates) despite the existence of work process interdependencies. An initial "top down" analysis during a re-design process may also provide useful criteria that can be further explored by using the Reconfig tool. In

[5]In the project examples provided here, we did not take into consideration criticality, but we recently improved the algorithm so that it is possible to distinguish between levels of criticality in future projects.

particular, by identifying the activities or business processes that are strategically important for an organization, the Reconfig tool can be set up to map precisely these activities and provide a map of interdependencies that reflect the activities that are most critical for the future of the organization.

Further development

As described, there are several limitations with our approach. However, when applying the Reconfig tool—or any tool for that matter—we do not believe that the relevant comparison is an ideal, theoretical model. What we should ask ourselves is whether the tool can improve current practice. So the most relevant question is whether the tool can serve as a decision making aid and improve the quality of decisions. As we noted in the introduction, the evidence that we have is that people's intuitive capacity for making grouping decisions is quite limited, and we would thus expect that an algorithm will outperform human decision makers.

Some of the limitations may also be addressed by further development. We have defined a research program to test and validate the algorithm that we use (Yu et al. 2007). We also believe that it should be possible to establish critical limits for key organization design variables. One approach is empirical. One can collect data by means of Reconfig and correlate the results with performance measures. One may use simulation tools such as SimVision (Levitt et al. 1999) to test the effects of alternative groupings of roles or sub-units on performance. One may also consider whether it is possible to borrow metrics from other fields to determine critical limits. As an example, in the field of fluid dynamics (Bradshaw 1970), one uses a metric called Reynolds number (Reynolds 1883), which represents the balance between chaotic and orderly processes. We are considering whether this metric can be used to assess the data collected by Reconfig and evaluate whether the coordination load of an organization is too low or too high.

Conclusion

We have described a new organization design tool—Reconfig—that is intended to support organization design decisions. It essentially operationalizes well-established principles related to grouping (March and Simon 1958; Thompson 1967) and modularity (Clark and Baldwin 2000; Sanchez 1995). The tool uses data collected by means of a survey questionnaire. The questionnaire is distributed to managers and employees in the organization and contains questions about their tasks/activities and working relationships. The tool takes these data as the input and applies a genetic clustering algorithm to identify a grouping (e.g., of teams, departments, etc.) that minimizes coordination costs. This can be the starting point for defining the formal structure of the organization.

We described two pilot applications of the tool, one in a university and one in a small manufacturing firm. The pilot applications have provided us with proof of concept. The algorithm identifies solutions that are deemed superior to manually developed solutions. At the same time, there are several limitations of the current version of the tool. Our current work seeks to validate the algorithm and to improve the functionality of the tool. Our end goal is to provide a tool that will effectively support organization design decisions and thereby increase organizational effectiveness.

Acknowledgments
We wish to thank two master level students, Andrea Karlsen and Celine Gronvold, for allowing us to use data from their thesis in this article (their thesis is included in the references).

Authors' contributions
The first author wrote the final version the article, but both of the co-authors contributed to the text as well. All of the authors contribute to the development of the tool being described, but the coding is done by the third author. All authors read and approved the final manuscript.

Authors' information
Nicolay Worren is an associate professor at the Norwegian University of Life Sciences and co-founder of Reconfig. He is the author of "Organization Design: Simplifying complex systems" (Routledge 2018) and writes a blog: www.organizationdesign.net
Tore Christiansen is an independent consultant in Norway and co-founder of Reconfig.
Kim Soldal is head of Enterprise Architecture for Capgemini in Norway and cofounder of Reconfig.

Funding
We received a small grant from the Technology Transfer Office of the Norwegian University of Life Sciences to develop the tool described in the article.

Availability of data and materials
The raw data used in the analyses in this article have been archived in the DataverseNO repository, https://doi.org/10.18710/5MS1JK

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Economics and Business, Norwegian University of Life Sciences (NMBU), P.O. Box 5003, 1432 Ås, Norway.
[2]Reconfig A.S., Inkubator Ås, Postboks 19, 1431 Ås, Norway.

Received: 1 May 2019 Accepted: 22 January 2020
Published online: 23 March 2020

References
Bradshaw P (1970) Experimental Fluid Mechanics. New York: Pergamon Press.
Browning TR (2001) Applying the design structure matrix to system decomposition and integration problems: a review and new directions. IEEE Trans Eng Manag 48(3):292–306
Buluç A, Meyerhenke H, Safro I, Sanders P, Schulz C (2016) Recent advances in graph partitioning. In: Kliemann L, Sanders P (eds) Algorithm Engineering. Lecture Notes in Computer Science, vol 9220, pp 117–158
Clark KB, Baldwin CY (2000) Design rules: the power of modularity. MIT Press, Cambridge, Mass
Cross R, Gray P (2013) Where has the time gone? Addressing collaboration overload in a networked economy. Calif Manag Rev 56(1):50–66
Danilovic M, Sandkull B (2005) The use of dependence structure matrix and domain mapping matrix in managing uncertainty in multiple project situations. Int J Proj Manag 23:193–203
Economist Intelligence Unit. (2015). Taming Organisational Complexity
Eppinger SD, Browning TR (2012) Design structure matrix methods and applications. In *Engineering systems*. MIT Press, Cambridge, Mass
Fang C, Lee J, Schilling MA (2010) Balancing exploration and exploitation through structural design: the isolation of subgroups and organization learning. Organ Sci 21(3):625–642
Freeman LC (2004) The development of social network analysis: a study in the sociology of science. Empirical Press, Vancouver B.C.
Galbraith JR (1974) Organization design: an information processing view. Interfaces 4(3):28–36
Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci U S A 99(12): 7821–7826
Goldberg DE, David E (2009) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading, Massachusets
Hansen MT (1999) The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits. Adm Sci Q 44(1):82–111
Holland JH, John H (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT Press, Cambridge, Mass
Jaques E (1989) Requisite organization: the CEO's guide to creative structure and leadership. Cason Hall & Publishers, Arlington, VA
Jaques E (2002) The psychological foundation of managerial systems. In: Keynote address for the mid-winter conference of the consulting psychology division (XIII) of the American Psychological Association (san Antonio, February 8, 2002)
Karlsen A, Gronvold C (2019) The design of organizations: how can one identify opportunities for improvements in small, fast growing firms by mapping interdependencies? Unpublished master thesis, Department of Economics and Business, Norwegian University of Life Sciences, Norway
Kilmann RH (1983) The costs of organization structure: dispelling the myths of independent divisions and organization-wide decision making. Acc Organ Soc 8(4):341–357

Levitt RE, Thomsen J, Christiansen TR, Kunz JC, Jin Y, Nass C (1999) Simulating project work processes and organizations: toward a micro-contingency theory of organizational design. Manag Sci 45(11):1479–1495
Livijn M (2019) Navigating in a hierarchy: how middle managers adapt macro design. J. Organ. Des 8(1):1–27
March JG, Simon HA (1958) Organizations. Wiley, New York
Mintzberg H (1983) Structure in fives: designing effective organizations. Prentice-Hall, Englewood Cliffs, N.J
Nagappan N, Murphy B, Basili V (2008) The influence of organizational structure on software quality: an empirical case study. In: Proceedings of the 30th international conference on software engineering, pp 521–530
Nuñez AN, Giachetti RE, Boria G (2009) Quantifying coordination work as a function of the task uncertainty and interdependence. J Enterp Inf Manag 22(3):361–376
Nutt PC (1999) Surprising but true: half the decisions in organizations fail. Acad. Manage. Exec 13(4):75–90
Reynolds O (1883) XXIX. An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels. Philos Trans R Soc Lond 174(1):935–982
Sanchez R (1995) Strategic flexibility in product competition. Strateg Manag J 16(S1):135–159
Schaeffer SE (2007) Graph clustering. Comp. Sci. Rev 1(1):27–64
Sharman DM, Yassine AA (2004) Characterizing complex product architectures. Syst Eng 7(1):35–60
Sherman JD, Keller RT (2011) Suboptimal assessment of Interunit task interdependence: modes of integration and information processing for coordination performance. Organ Sci 22(1):245–261 Retrieved from http://www.jstor.org/stable/20868857
Soldal KV (2012) Modularity as a solution to spatial interference in neural networks. (Unpublished master thesis, Department of Computer Science, Norwegian University of Science and Technology) Retrieved from https://brage.bibsys.no/xmlui/handle/11250/252948?locale-attribute=en
Sosa ME, Eppinger SD, Rowles C (2004) The coupling of product architecture and organizational structure in complex product development. Manag Sci 50(12):1674–1689
Steward DV (2007) Multiple heads acting as one make a thinking Enterprise. In: 6th IEEE international conference on cognitive informatics, pp 242–245
Thompson JD (1967) Organizations in action: social science bases of administrative theory. McGraw-Hill, New York
Van de Ven AH, Delbecq AL, Koenig R (2006) Determinants of coordination modes within organizations. Am Sociol Rev 41(2): 322–338
Watts DJ (2017) Should social science be more solution-oriented? Nat Hum Behav 1(1):1–5
Worren N (2018) Organization design: simplifying complex systems, 2nd edn. Routledge, Abingdon
Worren N, Cammelli F, Christoffersen K (2020) Organization design decisions: a card sorting approach. Working paper, School of Economics and Business, Norwegian University of Life Sciences, Aas, Norway
Worren N, van Bree J, Zybach W (2019) Organization design challenges: results from a practitioner survey. J. Organ Des 8(13): 1–18
Xu D, Tian Y (2015) A comprehensive survey of clustering algorithms. Annals of Data Science 2(2):165–193
Yu T-L, Yassine AA, Goldberg DE (2007) An information theoretic method for developing modular architectures using genetic algorithms. Res Eng Des 18(2):91–109

Publisher's Note

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

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com