Aalto University

Peltokorpi, Jaakko; Niemi, Esko

Effects of group size and learning on manual assembly performance: an experimental study

Effects of group size and learning on manual assembly performance: an experimental study

Jaakko Peltokorpi*

jaakko.peltokorpi@aalto.fi

Department of Mechanical Engineering, Aalto University School of Engineering, Espoo, Finland

Esko Niemi

esko.niemi@aalto.fi

Department of Mechanical Engineering, Aalto University School of Engineering, Espoo, Finland

*Corresponding author: Jaakko Peltokorpi, jaakko.peltokorpi@aalto.fi, +358 50 566 2382, Address: Puumiehenkuja 3, 02150 Espoo, Finland

Effects of group size and learning on manual assembly performance: an experimental study

In variable manual assembly production of highly customised products, effective allocation of workers to products is required. To support decision making here, industrial managers should be aware of the performance effects of the number of workers and learning within individual products. Evidence on such fundamental effects requires laboratory studies with products similar to those in real assembly industries. Because of the lack of such studies, this paper studies experimentally the effects of group size (one to four workers) and learning (up to four repetitions per group) on the performance of product assembly. The product, built for the purpose of the present study, consists of representative elements from real products in the mechanical engineering industry. A total of 68 undergraduate students participated in the experiments. The results from the experiments are in line with the hypotheses that the mean assembly time decreases at a decelerating rate as a function of both group size and repetitions, and that productivity per worker decreases as a function of group size. The results are explained in more detail through the experiences of the participants. Managerial implications and aspects for future research are also discussed.

Keywords: group size, learning, manual assembly, productivity, experimental study, mechanical engineering

1. Introduction

The assembly workforce is a crucial part of performance and costs in the traditional mechanical engineering industry. While assembly production faces variation as a result of e.g. volatile customer demand and differences between workers (Buzacott 2013), the ability to match the current capacity of the workforce with the floor-level processes is becoming increasingly important (Ruiz-Torres and Mahmoodi 2007; Saadat et al. 2013). In order to e.g. reduce labour costs or respond to schedule pressures, manning levels for certain products may become lower or higher than optimal (e.g. Sengupta and Jacobs 2004; Niemi 2009). On the evidence gained from the authors' experience, companies tend to have some degree of understanding of how worker assignments impact on assembly performance. However, e.g. production history data on worker assignments with the working hours spent is typically an inaccurate way of getting deeper knowledge of the effects of the size of the work group on the product. Another method is a field study, which uses interviews and direct observation in a natural setting (Krishnaswamy, Sivakumar, and Mathirajan 2009). It is, however, an interpretive method and hence not appropriate for the present study, which investigates the fundamental effects of group size and learning. As usual, controlled experiments in industrial conditions are practically impossible to perform. So far, some experimental studies on the effects of group size in manual tasks have been performed, such as those by Wicker et al. (1976), Sando, Tory, and Irani (2011), and Kuhrmann and Münch (2016). However, the tasks in those studies

2

are relatively small-scale and are clearly different from the tasks in real organisations. Experimental studies frequently use very simple tasks in which the participants have little, if any interdependence (Worchel, Wood, and Simpson 1992). Thus, there is a need for experimental studies that explore the more complex effects of group size with products similar to those in the real-life mechanical engineering industry.

This paper responds to the above-mentioned research gap by experimentally studying the effects of group size on assembly performance. Since permanent effects cannot be seen until the workers are sufficiently skilled, the assembly task needs to be repeated. This again makes it possible to examine the effects of learning with different group sizes. Effective learning is essential in the modern industrial production of variable products (Uzumeri and Nembhard 1998). The case product, built for the purpose of the present study, consists of representative elements from real products in the mechanical engineering industry. The subtasks and their interdependencies, as well as the size of the product, were designed in a way that enables groups of different sizes to coordinate different activities in several ways.

The participants in the experiments were 68 undergraduate students with no prior experience of the assembly product in question. The students were randomly assigned to groups, the sizes of which varied from one to four. Depending on the group, the assembly task was repeated up to four times in order to examine the effects of learning. In each trial, performance was measured as assembly time and productivity per worker. For these effects, hypotheses based on the literature were set. Additional information on the results was obtained from the experiences of the participants.

The rest of the paper is organised as follows. First, the related literature is reviewed. Second, the methods used for studying the effects of group size and learning on manual assembly performance are presented. Third, the results of the experiments are presented and discussed, and managerial implications are suggested. Finally, conclusions are drawn and some guidelines for further research are proposed.

2. Literature review

This section reviews the related literature to study the effects of group size and learning on manual assembly performance. More precisely, first, differences between workers, group performance with different types of tasks, and the effects of group size are reviewed. Second, the learning phenomenon and different types of learning curve models are reviewed.

2.1 Worker differences, group performance, and group size

Differences between workers are one of the major sources of variation and can cause great losses in production. Differences in individual ability reflect more on the variation of performance with increasing difficulty of tasks (Nembhard and Osothsilp 2002). In addition to this, differences between workers are revealed in a group working for a collective output. Buzacott (2013) gives an example related to a group of workers

3

performing at one station of an automotive assembly line. Even if the standard times for the tasks assigned to each worker were almost equal, the variability within and between the workers resulted in uncertain task completion times. Through time studies it was found that, for example, in a group of six workers, the best one typically performed twice as fast as the worst one. According to Dar-El (2013), this reflects the natural differences between workers. Steiner (1972) stated that when groups are assembled randomly from a population in which an attribute is normally distributed, the difference between the best and worst workers increases when the group size increases. In order to dynamically balance the assembly workload and tasks between workers, self-managed teams are valuable (Buzacott 2013). Such teams are an effective way to organise groups of workers for specific products and processes (Slomp and Molleman 2002).

The utilisation of the individual workers in a group depends significantly on whether and how the entire process can be divided into subtasks (in fact, the organisation of subtasks for teams and workers was already applied about a century ago by Henry Ford (Whitney 2004)). If a task is divisible, each subtask can be assigned to those who are the best for that particular subtask (Steiner 1972). The task type also determines the order in which different subtasks can be completed, and how, as well as the degree of worker collaboration necessary (Patel, Pettitt, and Wilson 2012). Some tasks require more coordination and communication between the workers. For this reason, groups with e.g. larger task loads may perform less efficiently (Qin, Nembhard, and Barnes 2015). In high-performing groups, the members are able to provide flexible assistance to overloaded colleagues (see Bukchin and Cohen (2013) in the context of work sharing). In such groups, the members have an awareness of task status, conditions, and roles (Buzacott 2013).

The following arguments are based on Steiner (1972). The complexity related to the group process (i.e. who does what, when, and how) clearly depends on the group size. When resources are used most advantageously, the level of potential productivity, i.e. the potential output per unit of time, for a group is reached. When this cannot be reached because of a faulty process (i.e. process loss), the actual productivity falls below the potential. For an individual working alone, this may happen as a result of performing steps that are necessary but are beneath his or her level of ability or performing them in an unfavourable order (see the paper by Lim and Hoffmann (2015), in which the order of assembly is considered). For two workers, the issues related to a single worker remain, in addition to which questions about the temporal and spatial coordination of the workers become relevant. When the group size increases, the number of coordination links between the workers increases rapidly (Steiner 1972).

Steiner (1972) makes a hypothetical statement (Figure 1) that is valid for many divisible tasks, according to which the potential productivity of a group increases at a decelerating rate and process losses at an accelerating rate as a function of group size. Thus, the actual productivity of a group first increases at a decelerating rate with group size and reaches its maximum at a certain group size, after which it decreases. This also means that the mean actual productivity per worker decreases as a function of group size.
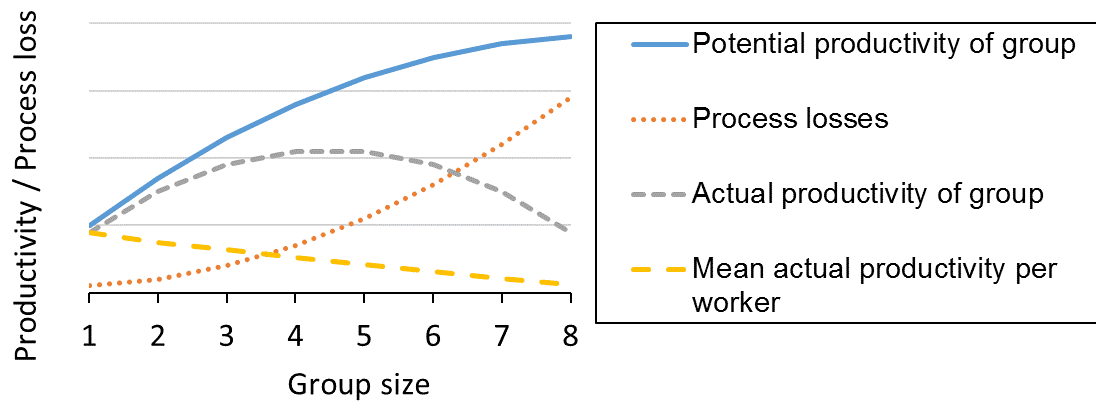
Figure 1. Hypothetical statement according to Steiner (1972).

Experimental research considering the effects of group size on performance is mainly carried out in the research field of social psychology (Thomas and Fink 1963; Wheelan 2009). Obviously, tasks in such research are often non-physical, such as generating ideas or solutions to problems (Frank and Anderson 1971) or answering quizzes (Littlepage 1991). When manual tasks have been considered, they have been relatively small-scale. For example, Wicker et al. (1976) studied performance in the context of racing a miniature slot car, Sando, Tory, and Irani (2011) in assembling small blocks, and Staats, Milkman, and Fox (2012) in assembling pieces of LEGO. Kuhrmann and Münch (2016) recently conducted an experimental study in which the dynamics of groups involving student participants were studied with a simple task. The objective of the task was to sort sweets, as far as possible, by colour and to write down the number of sweets sorted within a given time. The effects of group size and composition, stress factors, and exceptional situations were considered. The study showed that increasing the group size or re-forming existing groups resulted in more effort being needed for coordination and hence losses of productivity. In addition, when several new members joined a group that was already working, the working of the group was significantly disrupted. Additionally, there was not enough space for new members and some of them even ended up working outside the group's working space. Through repetition, the participants gained ideas and strategies relating to how to optimise the work.

Worchel, Wood, and Simpson (1992) criticise the fact that laboratory studies frequently use very simple tasks in which the participants have little, if any, interdependence. In real organisations, the tasks are more complex and the members of a group work closely with each other. Despite a lack of empirical evidence, some of the studies dealing with assembly systems assume that labour productivity (i.e. productivity per worker) decreases with an increasing number of collaborative workers. These studies, for example, aim to optimise the allocation of workers to products (Niemi 2009) or to compare the performance of worker coordination policies (Peltokorpi, Tokola, and Niemi 2015) or of cell and line assemblies (Sengupta and Jacobs 2004).

On the evidence of the literature review on the effects of group size, there is a need for research on tasks that are more complex and involve greater interdependence between workers. The research should be extended by considering tasks that have

elements from products in the assembly industry. With such tasks, broader-scale coordination of group members is needed between different subtasks.

2.2 Learning and learning curve models

Learning is central to enhancing productivity in manual assembly work, especially with high task complexity (Nembhard and Osothsilp 2002) and highly variable tasks (Uzumeri and Nembhard 1998), and when an experienced worker is replaced by an inexperienced one (Bukchin and Cohen 2013). In the literature, learning in labour-intensive tasks is typically associated with gaining experience through increasing cumulative production (Yelle 1979, Jaber 2011). Wright (1936) provided the earliest report on such a relationship in the form of a learning curve. This learning curve phenomenon relates to the observation that the direct labour hours used to produce a single unit decrease at a uniform rate as the number of units manufactured is doubled. The formula for this log-linear learning curve model is this:

$$T_x = T_1 X^n \tag{1}$$

where T_x denotes the time required to produce the Xth unit, T_1 the first unit, and X the cumulative unit number. $n = \log \phi / \log 2$ stands for the learning index, with ϕ denoting the learning rate, corresponding to the slope of the curve. Wright's model received criticism because by increasing X infinitely, the value T_x will eventually reach zero, which is practically impossible. De Jong (1957) then introduced what was termed a factor of incompressibility, denoted as M, which forces the learning curve to reach a steady-state level above zero (when $0 < M < 1$). The formula of this Plateau learning curve model is the following:

$$T_x = T_1 (M + (1 - M)X^n) \tag{2}$$

From the model, the steady-state time, denoted here as T_M, can be calculated as $T_M = T_1 * M$. De Jong (1957) suggests that when technical equipment and work organisation remain unchanged the learning curve will plateau. De Jong's model thus assumes two components in each task, one subject to improvement according to the learning rate and the other subject to no improvement (Jaber 2006). Conway and Schultz (1958) first observed the phenomenon of plateauing, the causes of which were later explained by several researchers. For example, plateauing may be associated with workers ceasing to learn or the lack of necessary technological improvements (Yelle 1979), depreciation of knowledge (Li and Rajagopalan 1998), or constant quality problems (Jaber and Guiffrida 2004). Consequently, performance improves through quality improvements occurring when inefficiencies in production are discovered and learning takes place (Fine 1986). A plateau barrier can also be broken by additional investments in training or new technology (Jaber and Guiffrida 2004). It is noteworthy that, for inexperienced workers, the learning

rate and steady-state level may vary. This variation increases with increasing task complexity (Nembhard and Osothsilp 2002).

Argote, Beckman, and Epple (1990) criticised the fact that the conventional learning curve models (such as those presented above) significantly overstate the persistence of learning. They showed that, rather than cumulative output, depreciation of knowledge resulting from breaks at work provides a better prediction of current production. Since then, several studies have considered forgetting with learning curve models that are also applicable to industrial settings (e.g. Jaber and Bonney 1996; Nembhard and Uzumeri 2000; Sikström and Jaber 2002). In addition to learning and forgetting, Jaber, Givi and Neumann (2013) incorporated worker fatigue and recovery into their model.

Dar-El, Ayas, and Gilad (1995) developed a learning model that combined cognitive and motor skills, which also means that learning would not occur at a constant speed, as is assumed with e.g. Wright's model. A review by Hill (1982) suggests that the individual capacity to learn and cognitive stimulation are underlying factors that also affect the group process. Schilling et al. (2003) studied the effects of task variation on learning in groups. They found that, compared to specialisation, learning was faster when working on different but similar types of problem-solving tasks. Jaber and Guiffrida (2004) modified Wright's learning curve model and developed a combination of two learning curves: the reduction of time is described for each additional unit produced (the first curve), and for each additional defective unit reworked (the second curve). Depending on learning rate with rework, the shape of the composite learning curve was convex or similar to that of De Jong (1957) or Wright (1936). Obviously, when the time required to rework a defective item becomes insignificant, the curve returns to a shape similar to that of Wright (1936).

Alongside learning curves for individuals, group learning curves have received growing attention. As reviewed in Section 2.1 above, group performance is dependent on a multitude of factors, many of which are associated with interaction among group members. For this reason too, group learning curves include a social component when knowledge is transferred among individuals in a group (e.g. Schilling et al. 2003; Wilson, Goodman, and Cronin. 2007; Ryu et al. 2005; Ingram and Simons 2002; Glock and Jaber 2014). Glock and Jaber (2014) developed a group learning curve model that follows the same form as that of Wright (1936) and includes the compatibility of knowledge, the willingness and ability of the group members to share and absorb knowledge, and the number of members in the group.

The literature review shows that learning is dependent on several factors and these factors have increasingly been incorporated into the learning curve models. Even if much progress has been made since the earliest model of Wright (1936), that model continues to be a good starting point for further developments of learning models. Jaber (2011) and Dar-El (2013) provide more comprehensive reviews of different learning curve models.

3. Methods

In order to study the effects of group size and learning on manual assembly performance, this section first sets the hypotheses for the experiments in the present study. Then the assembly product and the experimental design to test the hypotheses are described.

3.1 Hypotheses

For the experiments, this study sets three hypotheses. The first hypothesis is based on the Plateau learning curve model (De Jong 1957), which assumes that learning reaches a steady-state level above zero (a modification of Wright (1936)), after which no further improvements in performance occur. The plateauing may occur because the technical equipment and work organisation remain unchanged (De Jong 1957) or there are constant quality problems (Jaber and Guiffrida 2004). Plateauing is simply associated with workers ceasing to learn (Yelle 1979).

> [Hypothesis 1]: For each group size, the mean assembly time decreases at a decelerating rate as a function of repetitions. At a certain repetition, the mean assembly time reaches its steady-state level above zero.

The other two hypotheses are based on a hypothetical statement by Steiner (1972). According to this statement, the progress of work does not speed up, or assembly time does not decrease, in direct proportion to the number of additional workers and loss of labour productivity, i.e. productivity per worker, occurs.

> [Hypothesis 2]: For each repetition, the mean assembly time decreases at a decelerating rate as a function of group size.

> [Hypothesis 3]: For each repetition, the mean productivity per worker decreases as a function of group size.

In practice, productivity losses will occur because of the increasing complexity of the temporal and spatial coordination of the workers. According to the experiences of the present authors and industrial professionals, this phenomenon is typical of the industrial assembly work of relatively large products. To assemble such a product, one worker is typically required. Additional workers are then allocated to the product for multiple reasons, such as schedule tightness or low overall workload in production. The size of the product and the number and independence of subtasks makes it possible for several workers to assemble the product simultaneously.

3.2 Case product

The product used in the experiments is shown in Figure 2(a). This product was constructed for the purposes of the present study. It does not directly represent any real product from the assembly industry but it consists of representative elements, i.e. its

structure and parts, from real products. The parts (or subtasks) and their interdependencies, as well as the size of the product (approximate edge length of 1.5 and height of 2 metres), are designed in a way that enables groups of different sizes to coordinate different activities in several ways. This type of non-standardised coordination of workers is typical in companies that lack deeper knowledge of the effects of work group size on the product. In the present product, different types of parts (the pipe subassemblies (P), hoses (H), modules (M), plate (PL), and valve (V)) are assembled in a single frame made of steel. The precedence constraints of the parts are shown in Figure 2(b). The figure presents the most suitable assembly sequence (left to right) among the interconnected parts (subsystems). For these parts, the assembly processes can partly overlap. Different subsystems can be assembled in parallel. They can also be assembled separately from the assembly frame but this did not really happen as it would make the handling of the subsystems even more difficult.
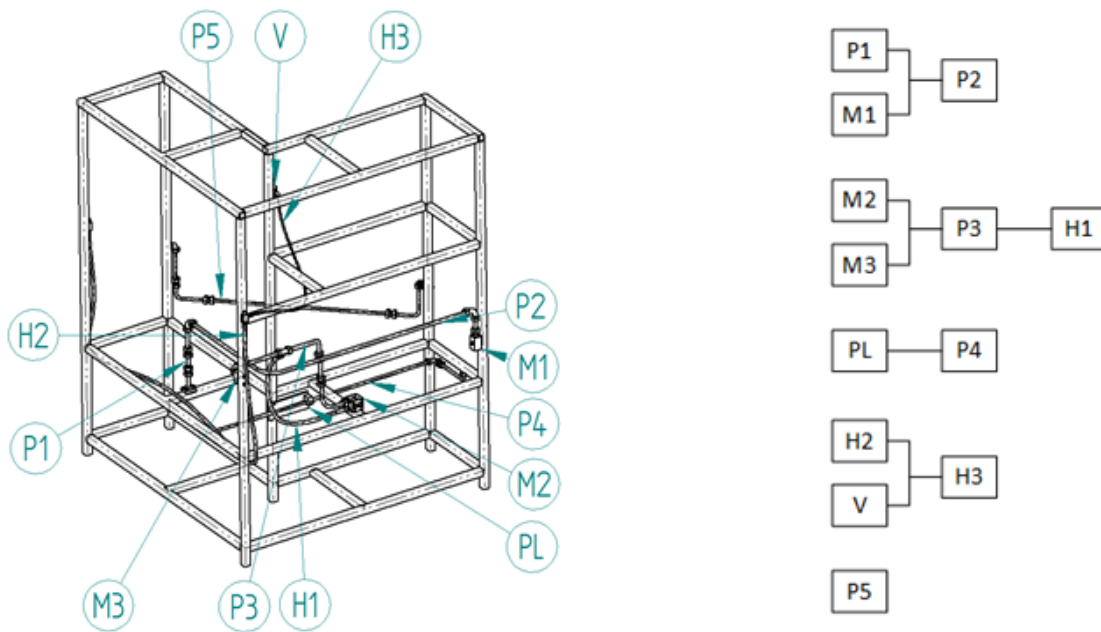
Figure 2. (a) The case assembly product. (b) Precedence constraints of parts.

In order to avoid excessively large work content and long assembly times, pre-tests for the assembly task were conducted utilising the laboratory staff. Finally, the total number of different parts was reduced to 13 (frame and small parts excluded), a number which is clearly less than is the case with typical products in industry. As Figure 2(a) shows, the mounting locations of the different parts are relatively evenly distributed around the product. This enables a work group to divide the subtasks between workers. With the large pipe parts (P2 and P5), worker cooperation is also justified. Concerning a suitable assembly sequence, possible precedence constraints (physical), together with different work contents (temporal) of subtasks, should be taken into account. All the above makes the coordination of a group, i.e. who does what, when, and how, important.

The main limitation with the case product is the fact that it does not have a functionality, as is required from real products. This further raises questions about quality requirements and the actual quality of a finished product. Despite the limitation regarding functionality, the finished case product and thus the assembly process need to meet certain predetermined quality requirements, the fulfilment of which was controlled in the experiments, as presented in the next section.

3.3 Experiments

The experiments were conducted in the Production Engineering Laboratory at Aalto University School of Engineering, Espoo, Finland, between 9 September and 2 October 2015. The experiments were carried out approximately between 12 noon and 6 p.m. A total of 68 male students participated in the experiments as a part of their mechanical engineering course at the third-year bachelor's level. In order to test the effect of group size, the participants performed the assembly task either alone ($N=9$) or in a group of two ($N=10$), three ($N=9$), or four ($N=3$) workers. Each participant was randomly assigned to one of a total of 31 groups. The low number of four-worker groups is due to the limited number of participants within the course in question. The students used in the experiments represent a sufficiently homogenous population. Using specialised assembly workers from a specific company is not appropriate for the present study as it would weaken the randomisation of the groups.

In order to test the effect of learning, the participants repeated the assembly task several times. The number of repetitions was first set to three, and then, on the basis of the performance of the first few groups, it was set to four. The final numbers of repetitions for each group size can be seen in Table 1. It is noteworthy that only two of the nine participants who worked alone repeated the task four times. This was due to the relatively large work content and obvious fatigue at the fourth repetition for single workers. Thus, three repetitions were considered as default for single workers.

The layout of the assembly cell is shown in Figure 3. The figure shows the positions of the assembly product (frame), parts table, and movable rack for the assembly drawing, small parts, and tools in the assembly cell. In addition, the main dimensions (in millimetres) of the layout are given. A video recorder was placed obliquely in order to get a general view of the assembly cell.

Figure 3. Assembly cell layout.

Before the first experiment, an information lecture regarding the purpose of the research and general instructions was given. The participants were asked to reserve time enough for the experiment that their other activities did not affect their assembly performance. The participants were informed personally about the time they should arrive for the experiment. This was to ensure that the group dynamics were not built until the experiment.

In the experiment, the participants were first introduced to the assembly task. This included an overview of the assembly drawing and the information it contained (Appendix 1). Information on using the tools was also given. This included the types of tools, the requirement always to use a tool for the final tightening of a bolt or screw, and that they should be tightened to a moderate torque. The parts table and screw boxes were also shown. The names and positions of the parts were not marked on the table, but before each trial, the parts were placed in approximately the same positions. The rack for the drawing, small parts, and tools was organised according to lean principles. This meant that only the necessary numbers of small parts and tools were given, with the proviso that the number of tools could not constrain the assembly process. The types and sizes of the bolts and screws were marked in their dedicated screw boxes. The position of each tool was marked on the rack. The wrench sizes were marked as numbers and hex key sizes as colour codes.

Regarding the assembly process itself, the instruction was given that it was the group's decision how they would complete the task. In addition, the experimenter stated that he would not in principle interfere in the assembly process. In practice, this meant

11

that only if a group had consumed an unreasonable amount of time and was not going forward with installing a part were small hints given on the experimenter's initiative. Typically, the hints were related to the positioning of a part. This principle of giving hints would be equivalent to a foreman's guidance in assembly work. The quality of the assembly product was only controlled in that if the product was incomplete or a tool had not been used for the final tightening the participants were notified. At the beginning of a trial, the participants were also told that the assembly time would be recorded but that they should work at a normal pace, without hurrying. This instruction about the pace of the assembly was given before each trial. After that the experimenter gave permission to start working and started the video and time recording.

In each repetition, when the last part was completely installed, the experimenter stopped the recording. The product was disassembled by the laboratory staff. At the same time, the participants had a break and answered questions related to the repetition that had just ended. The questionnaire is presented in Appendix 2. In the questions, the participants were asked to evaluate their attitudes to the processing pace they had been instructed to follow. In addition, questions were asked about their experiences of inefficiency in working as a group and of learning. This also made it possible for the participants to reflect more consciously on their work. After the last repetition, questions related to group size, assembly instructions, and how comfortable the participants felt about the test were asked. At the very end of the experiment, the participants were instructed not to say anything about the test event or the assembly product to those who had not yet participated in the experiment.

In order to study the effects of group size and learning on manual assembly performance, the following assumptions in the experiments hold. The assumptions are derived from the above-described procedures or are based on the literature or authors' view.

1. The workers (students) have no prior experience of the case assembly product and are not assembly professionals.
2. The workers have no prior experience of working in the group they are assigned to.
3. The physical and cognitive abilities may vary between the workers. By following the instructions, workers should have sufficient abilities to assemble the case product.
4. The experimenter's guidance with possible installation problems is equitable for the workers and will help the workers in the progression of the work.
5. Repetitions improve the motor and cognitive skills with the case assembly.
6. The workers and groups may reflect differently on their work during repetitions and breaks. This will also affect improvements in motor skills, assembly sequence and division of work, etc.
7. The mental stress caused by the test situation may affect the performance of the workers.

8. Fatigue is not caused significantly by the given number of repetitions and due to breaks between them.
9. Increased group size reduces the physical space per worker and makes communication among the workers more complicated.
10. The quality of the assembly parts and tools, and other external conditions in the experiments, remain unchanged.
11. The quality criteria for the finished products remain unchanged.

In the experiments, the performance of the assembly work is measured as assembly time and productivity per worker. For each group size, the mean assembly time (T_{mean}) for each repetition is calculated as the average of the assembly times T_i of different groups i $=1...N$:

$$T_{mean} = \frac{\Sigma_i^N T_i}{N} \qquad (3)$$

For each group size, the mean productivity per worker (P_{mean}) for each repetition is calculated as the inverse of T_{mean} divided by the group size (S):

$$P_{mean} = \frac{1}{T_{mean}*S} \qquad (4)$$

The mean productivity per worker corresponds to the harmonic mean, which is typically used for calculating mean rates.

4. Results and discussion

This section presents the results from the experiments. First, descriptive statistics to study assembly time and productivity per worker with different group sizes and repetitions are shown. Secondly, the results from the questionnaires are presented. Each result is presented and discussed.

4.1 Assembly time and productivity per worker

Table 1 shows descriptive statistics from the experiments. First, for each group size and repetition, the sample size (N) is shown. Then the statistics for assembly time and productivity per worker in terms of mean value (mean) and standard deviation (SD), as well as the minimum (min) and maximum (max) value, are presented. Generally, for all the results, the extremely small sample sizes for one-worker groups at the fourth repetition ($N=2$) and for four-worker groups ($N=3$ in each repetition) increase the uncertainty of the results.

Table 1. Descriptive statistics of experiments.

Group size		Assembly time (hours)					Productivity (products/hour/worker)				
		Repetition					Repetition				
		1	2	3	4	Tot	1	2	3	4	Tot
1	N	9	9	9	2	29	9	9	9	2	29
	Mean	0.694	0.419	0.346	0.327	0.475	1.440	2.386	2.892	3.059	2.103
	SD	0.114	0.050	0.046	0.053	0.169	0.233	0.258	0.354	0.507	0.690
	Min	0.563	0.366	0.279	0.289	0.279	1.107	1.887	2.214	2.742	1.107
	Max	0.903	0.530	0.452	0.365	0.903	1.776	2.733	3.582	3.458	3.582
2	N	10	10	10	7	37	10	10	10	7	37
	Mean	0.419	0.239	0.200	0.177	0.265	1.192	2.095	2.506	2.824	1.885
	SD	0.144	0.079	0.056	0.043	0.132	0.343	0.740	0.705	0.665	0.878
	Min	0.254	0.141	0.134	0.125	0.125	0.633	1.230	1.627	1.985	0.633
	Max	0.790	0.407	0.307	0.252	0.790	1.965	3.536	3.742	4.009	4.009
3	N	9	9	9	7	34	9	9	9	7	34
	Mean	0.345	0.189	0.131	0.130	0.203	0.967	1.759	2.535	2.569	1.643
	SD	0.077	0.020	0.015	0.026	0.099	0.260	0.197	0.270	0.460	0.729
	Min	0.223	0.159	0.115	0.110	0.110	0.735	1.565	2.091	1.920	0.735
	Max	0.453	0.213	0.159	0.174	0.453	1.493	2.091	2.892	3.038	3.038
4	N	3	3	3	3	12	3	3	3	3	12
	Mean	0.287	0.137	0.119	0.106	0.162	0.871	1.829	2.095	2.348	1.539
	SD	0.009	0.015	0.018	0.021	0.077	0.026	0.198	0.334	0.525	0.668
	Min	0.279	0.123	0.101	0.083	0.083	0.843	1.636	1.826	2.079	0.843
	Max	0.296	0.153	0.137	0.120	0.296	0.896	2.032	2.486	3.020	3.020
Tot	N	31	31	31	19	112	31	31	31	19	112

In order to test whether the mean assembly times in Table 1 follow the Plateau learning model (Equation 2), the models were fitted to measured data using the least squares method. Figure 4 presents the mean assembly time of each group size (workers, w) in relation to repetitions and the Plateau models fitted to the data. The figure also presents the mean time decrease (%) with group sizes of two, three, and four in relation to a group size of one for each repetition.
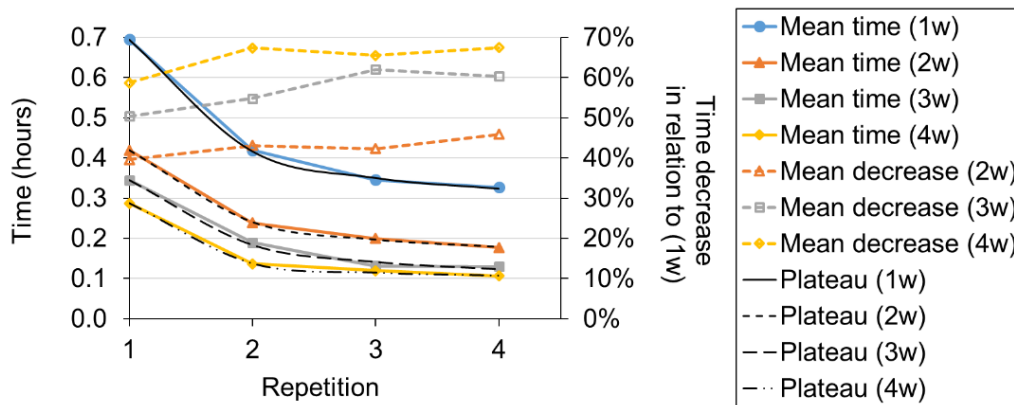

Figure 4. Mean assembly time of each group size (workers, w) in relation to repetitions and the Plateau models fitted to data. Decrease in mean times (%) with group sizes of two, three, and four in relation to a group size of one for each repetition.

As Figure 4 shows, for each group size, the mean assembly time decreases at a decelerating rate as a function of repetitions. Thus, the first part of Hypothesis 1 is

validated. Clearly, the Plateau model (De Jong 1957) fits the data almost perfectly. Obviously, if the number of repetitions increases, the assembly time soon reaches the steady-state level above zero. The Plateau model shows the learning rate $\phi = 0.33$, factor of incompressibility $M = 0.40$, and the corresponding steady-state assembly time $T_M = 0.277$ hours for one-worker groups. For two-worker groups, $\phi = 0.34$, $M = 0.35$, $T_M = 0.147$, for three-worker groups, $\phi = 0.38$, $M = 0.25$, $T_M = 0.087$, and for four-worker groups, $\phi = 0.20$, $M = 0.35$, $T_M = 0.101$. The slight increase in T_M when the group size is changed from three to four needs more detailed investigation in a future study. One explanation for this exception is the uncertainty of the results as a result of the small sample size ($N=3$) with four-worker groups. The calculated T_M is sensitive to changes in measured data.

The mean assembly time decreases at a decelerating rate as a function of group size with each repetition, except with Repetition 2 for a group size of four. This is in line with Hypothesis 2 and indicates that with a certain group size, the mean assembly time reaches a minimum. In relation to a one-worker group, the approximate decrease with two workers is 40-46%, with three workers 50-60%, and with four workers 59-67%, depending on the repetition. The slightly increasing curves of relative difference indicate that, in relation to a one-worker group, larger groups utilise their potential better with an increasing number of repetitions.

Whether the assembly time differs significantly depending on the group size or number of repetitions was examined using Pearson's chi-squared test. This test was chosen because of the fact that with the relatively low sample sizes, the normality of the data in all the different treatments could not be validated (which is presumed e.g. with ANOVA). In order to run a chi-squared test, first, the recorded assembly times were divided into four simple categories (less than 0.25; 0.25-0.5; 0.5-0.75, and more than 0.75 hours). Then, for each group size and repetition, the frequencies in each time category were calculated. By continuing the procedure it could be determined whether the frequency counts are distributed identically across different populations. At low p-values ($p < 0.05$) the counts are non-identically distributed, and thus the response (assembly time) differs significantly, depending on group size or repetition. As a result, for each repetition (1-4), the effect of group size on assembly time is significant ($p < 0.006$). For each group size (1-4), the effect of repetition is also significant ($p < 0.007$).

Figure 5 presents the mean productivity per worker of each group size (workers, w) in relation to repetitions. The figure also presents the mean productivity loss (%) with group sizes of two, three, and four in relation to a group size of one for each repetition.
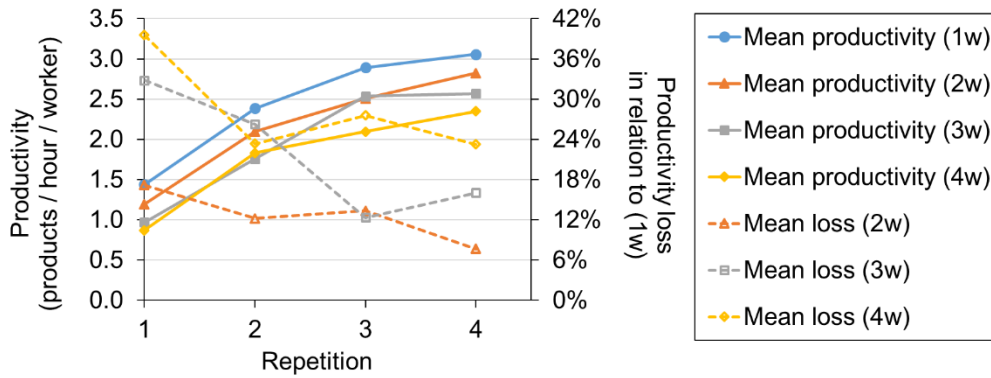
Figure 5. Mean productivity per worker of each group size (workers, w) in relation to repetitions. Mean productivity loss (%) with group sizes of two, three, and four in relation to a group size of one for each repetition.

As Figure 5 shows, for each group size, the mean productivity per worker increases as a function of repetition. Across the different group sizes, on average, an 83% increase in productivity is reached between the first and second repetition, and the third repetition increases it by 25%. By repeating the task twice, on average, 57% of the total increase in productivity is reached. These can be characterised as huge improvements in performance, reflecting fast learning.

In relation to a one-worker group, the productivity loss with two workers is approximately 8-17%, with three workers 12-33%, and with four workers 23-40%, depending on the repetition. On average, the productivity loss decreases as a function of repetition. This again shows that groups (of at least two workers) can utilise their potential productivity better as the number of repetitions increases. It is noteworthy, as the standard deviations in Table 1 show, that the between-group difference in productivity per worker is greatest with two-worker groups. Additionally, the relative standard deviation (the coefficient of variation) is greatest with two workers (not shown in the results).

Figure 6 presents the mean productivity per worker in relation to group size for each repetition (Rep).
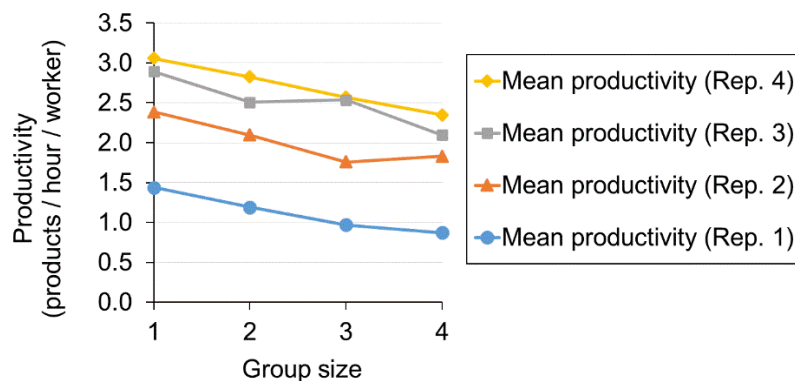

Figure 6. Mean productivity per worker in relation to group size for each repetition (Rep).

As Figure 6 shows, the mean productivity per worker decreases quite steadily as a function of group size. Two exceptional cases can be observed: for Rep. 2 between group sizes of three and four, and for Rep. 3 between group sizes of two and three, the mean productivity per worker increases slightly as the group size increases. The first exception is, however, uncertain because of the small sample size ($N=3$) with four-worker groups. For the second exception, a simple, but not verified, explanation would be that, in larger groups, the workers had learned the ways of working through specialisation faster by Rep 3. To find out a practical explanation of this, additional experiments would be needed. In general, Hypothesis 3 can be validated and thus the results are in line with the statement by Steiner (1972).

Whether productivity differs significantly depending on group size or repetition was again examined using Pearson's chi-squared test. Productivity was divided into four simple categories (less than 1; 1-2; 2-3 and more than 3 products per hour per worker). For each repetition (group size), the significance of group size (repetition) was tested not only across all group sizes (repetitions) but also pairwise. The resulting p-values are presented in Table 2. A grey background indicates that the effect is significant ($p < 0.05$).

Table 2. Significance of the effects of group size and repetition on productivity per worker.

(a) Significance of the effect of group size on productivity for each repetition.

Rep. 1

	2	3	4
1	0.330	0.003	0.001
2	-	0.011	0.003
3	-	-	0.248
Total			0.001

Rep. 2

	2	3	4
1	0.030	0.004	0.054
2	-	0.137	0.532
3	-	-	0.700
Total			0.011

Rep. 3

	2	3	4
1	0.179	0.134	0.157
2	-	0.019	0.420
3	-	-	0.070
Total			0.071

Rep. 4

	2	3	4
1	0.773	0.773	0.709
2	-	0.513	0.490
3	-	-	0.788
Total			0.882

(b) Significance of the effect of repetition on productivity for each group size.

Group size = 1

	2	3	4
1	0.000	0.000	0.004
2	-	0.216	0.081
3	-	-	0.425
Total			0.000

Group size = 2

	2	3	4
1	0.030	0.004	0.004
2	-	0.621	0.435
3	-	-	0.784
Total			0.024

Group size = 3

	2	3	4
1	0.008	0.000	0.007
2	-	0.001	0.030
3	-	-	0.093
Total			0.000

Group size = 4

	2	3	4
1	0.050	0.050	0.050
2	-	0.414	0.189
3	-	-	0.368
Total			0.037

As Table 2(a) shows, the effect of group size is less significant when the assembly task is repeated. In general, the effect of repetition [Table 2(b)] on productivity is more significant than the effect of group size. These results show that learning through repetition is more critical to performance than group size. Generally, in all results, the extremely small sample sizes for one-worker groups at the fourth repetition ($N=2$) and for four-worker groups ($N=3$ at each repetition) increase the uncertainty of the results.

4.2 Questionnaires

The participants felt that the instructions were clear (mean 4.57 on a Likert scale of 1-5, Question 5, Q5, in Appendix 2) and the exercise was comfortable (mean 4.60, Q6). The participants also endeavoured to follow the instructions to work at a normal pace without hurrying (mean 4.58 of all trials, Q1). Thus, on the evidence of the opinions of the

participants, the test situation itself does not seem to have had a significant impact on the results. Table 3 presents the descriptive statistics for the extent of the learning that was experienced in relation to previous repetition (Q3) and of the inefficiency that was experienced in group performance (Q2).

Table 3. Descriptive statistics for learning experienced and inefficiency experienced for each group size in different repetitions (Likert scale of 1-5, 1=not at all, 2=rather little, 3=neither much nor little, 4=quite much, 5=very much).

Group size		Learning experienced				Inefficiency experienced				
		Repetition				Repetition				
		2	3	4	Tot	1	2	3	4	Tot
1	N	9	9	2	20					
	Mean	4.11	3.22	3.00	3.60	-	-	-	-	-
	SD	0.60	0.97	1.41	0.94					
2	N	20	20	14	54	20	20	20	14	74
	Mean	4.15	2.95	2.29	3.22	1.30	1.10	1.15	1.14	1.18
	SD	0.75	0.76	0.73	1.06	0.73	0.31	0.37	0.36	0.48
3	N	27	27	21	75	27	27	27	21	102
	Mean	4.15	3.19	2.38	3.31	1.96	1.48	1.33	1.48	1.57
	SD	0.72	1.11	0.92	1.16	0.85	0.51	0.55	0.60	0.68
4	N	12	12	12	36	12	12	12	12	48
	Mean	4.25	3.33	2.50	3.36	2.17	2.08	1.92	2.25	2.10
	SD	0.45	1.30	1.24	1.27	0.83	1.00	1.00	1.22	0.99
Tot	N	68	68	49	185	59	59	59	47	224

Figure 7 illustrates (a) the mean learning that was experienced and (b) the mean inefficiency that was experienced in group performance with each group size for each repetition.
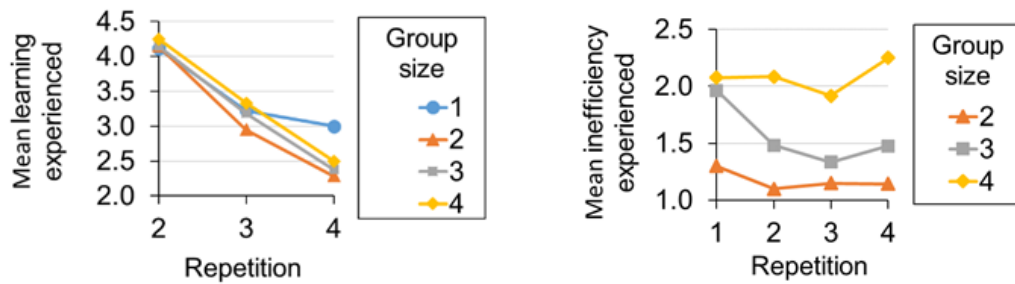

Figure 7. (a) Mean learning experienced. (b) Mean inefficiency experienced.

As Figure 7(a) shows, unexpectedly, the experience of learning decreases as a function of repetitions. The effect of repetition (across all repetitions) on learning that was experienced is significant for two- ($p < 0.000$), three- ($p < 0.000$), and four-worker ($p < 0.014$) groups and non-significant for those that worked alone ($p < 0.391$). For each repetition, the effect of group size on learning that was experienced is non-significant (for Rep.2, $p < 0.972$; for Rep.3, $p < 0.34$, and, for Rep. 4, $p < 0.551$). Thus, in general, learning was experienced similarly, irrespective of group size.

As Figure 7(b) shows, the experience of inefficiency in working as a group increases as a function of group size. In general, the extent of the inefficiency that was experienced is relatively low: across all repetitions, on average, 1.18 for two-, 1.57 for three-, and 2.1 for four-worker groups (Table 3). Thus, most workers in two-worker groups did not experience inefficiency at all, and in four-worker groups rather little. With these group sizes the experience seems to remain relatively unchanged across different repetitions and the effect of repetition is non-significant for the two-worker ($p < 0.8$) and for the four-worker groups ($p < 0.771$)), whereas with the three-worker groups ($p < 0.063$), changes are clearly visible. For each repetition (except for the third) the effect of group size on the inefficiency that was experienced is significant (for Rep. 1, $p < 0.009$, for Rep. 2, $p < 0.000$, for Rep. 3, $p < 0.077$, and, for Rep. 4, $p < 0.024$). Thus, in general, the experience of inefficiency depends on group size.

Figure 8(a) presents the mean experience of the suitability of the size of the current group in relation to the size of the current group (on a Likert scale of 1-3, 1=too small, 2=appropriate, 3=too large, Q4 in Appendix 2). Figure 8(b) presents the mean opinion on the appropriate group size (Q4) as a function of the size of the current group.
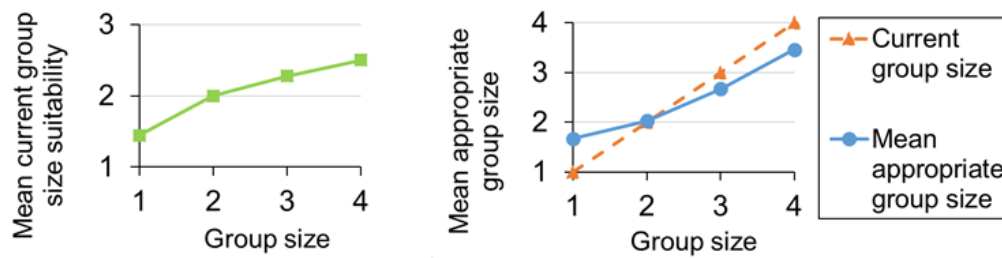

Figure 8. (a) Mean suitability of current group size and (b) mean appropriate group size in relation to current group size.

As Figures 8(a) and 8(b) show, all the participants in the two-worker groups were content with the size of their group. According to the answers to the open question (Q4), they found the size of their group appropriate, because one can help another, especially with large and heavier pipe parts. They also mentioned that a third worker would not bring added value, mainly because of the limitations of the physical space. Almost half (four out of nine) of those who worked alone felt the size of their group was suitable. However, three of them found 1.5 an appropriate group size (meaning a temporary presence and help from another worker). Typically, help from another worker would have been needed with large pipe parts. Almost three out of four participants in the three-worker groups felt the size of their group was suitable. Typically, they felt that there were enough tasks for three workers. While two workers are cooperating on one part, the third can focus on another part. On the other hand, some felt that after some routine in the task had been established, two workers would be optimal.

Surprisingly, half of the participants in the four-worker groups felt the size of their group was suitable. Some mentioned that three workers would be better because of the physical space limitations (an experience of congestion similar to that in Kuhrmann and

Münch 2016). Some also felt that there was not enough work for all four workers at the end of the task (i.e. an unequal workload between the workers). It is noteworthy that nobody from the four-worker groups found two workers to be an appropriate group size. The opinions of the participants in larger groups indicate that within a relatively small number of repetitions (up to four), these participants felt they learned to work as a group but did not see too much harm in the group size. Thus, when estimating an appropriate group size from the performance point of view, participants in larger groups seem to pay relatively little attention to process losses (as a result of e.g. more complicated coordination of workers). Underestimation of declining efficiency with larger groups is also supported in the experimental study by Staats, Milkman, and Fox (2012). In their study, unlike in the present study, estimations were asked for before completing the task and in terms of the total number of minutes it takes a group to complete a task.

In the present study, the experience of the suitability of the size of the current group differs significantly between those that worked alone and those in two-worker groups ($p < 0.000$) and between the two- and three-worker groups ($p < 0.012$). Between the three- and four-worker groups, the difference is non-significant ($p < 0.163$). These results indicate that smaller groups are most critical with regard to the view on the suitability of the group size. If the group size is increased, the differences between these opinions will disappear. The opinions on appropriate group size differ significantly between each adjacent group size ($p < 0.004$ between the one- and two-worker, $p < 0.000$ between the two- and three-worker, and, $p < 0.001$ between the three- and four-worker groups). Thus, one's own group size clearly affects the worker's view on what is an appropriate group size.

5. Managerial implications and limitations of the research

This study contributes to the understanding of the effects of group size and learning with industrial assembly products. Such products are complex and often involve great interdependence between workers. This study has shown a number of significant effects and aspects that have managerial implications and possible limitations:

(1) The results from the experiments give evidence on the effects of group size and learning on manual assembly performance. The product used in this study has, more than previous studies (Wicker et al. 1976; Sando, Tory, and Irani 2011; Staats, Milkman, and Fox 2012; Kuhrmann and Münch 2016), elements from real assembly products, which makes the research environment and the effects that were studied more realistic. The experimental design proposed in this paper can be customised to different products and conditions and will help in optimising group size with respect to e.g. minimum assembly time or maximum productivity per worker.

Limitation: in practice, controlled experiments in fully industrial conditions are impossible to perform. Also, in real conditions, optimal group sizes may vary depending on the stage of the assembly work, and thus one needs to investigate these stages separately from each other.

(2) On the evidence of the opinions of the participants and observations made in the experiments, several characteristics of the subtasks affect a suitable group size for a product. The number of subtasks determines the maximum number of workers they can be divided between, the distance between the subtasks whether the workers have space enough to work, the independence of the subtasks whether the workers can assemble them in parallel or in sequence, and the physicality of the subtasks whether another worker is needed to help. All the above characteristics also influence how a given number of workers can be coordinated between subtasks.

Limitation: in the real-life assembly industry, decisions on worker assignments are more dynamic and also based on external factors, such as the total number of workers available, demand, schedule tightness, and various disturbances (Saadat et al. 2013). Additionally, factors related to individuals, such as skills (De Bruecker et al. 2015) and chemistries between workers, have their own effects.

(3) The learning curves within the case product for novice workers and groups embody two noteworthy things. First, the learning is very fast, and the increase in productivity is already tremendous when the task is repeated twice. This is important when workers frequently have to master new tasks (see Uzumeri and Nembhard 1998), which is typical in the small batch production of highly customised products. Second, plateauing occurs quite rapidly and the steady-state level in performance is almost reached. One probable cause of plateauing in this study is unchanged work organisation and technical equipment (see De Jong 1957). This highlights the importance of managers providing improvements in these matters to make further learning possible. In general, the learning effects with groups of different sizes contribute to the research on group learning curves (see Glock and Jaber 2014), whose applicability for industrial settings needs a lot of further research.

Limitation: in contrast to the present study, the productivity increase is lower in industrial conditions (for typical progress ratios, see Badiru 2005, p. 539), where workers normally possess prior experience because of their professional background and experience of similar product variants. Another aspect is that one may become familiar with the case product in a short time as it has fewer quality requirements and complexity compared to real products.

(4) Even if the workers in larger groups experienced, to some extent, inefficiency in group working, from the performance point of view, they were surprisingly happy with their group size. Such an underestimation of declining efficiency with larger groups (see also Staats, Milkman, and Fox 2012) needs attention from industrial managers.

6. Conclusions

This paper performed an experimental study of the effects of group size (one to four workers) and learning (up to four repetitions per group) on manual assembly performance. The case product, which consisted of representative elements from real products in the mechanical engineering industry, was designed in a way that enabled the groups to coordinate different activities in several ways. The results from the experiments show the performance as the mean assembly time and mean productivity per worker. Additional information on the performance was obtained from the experiences of the participants.

In general, the results validated the hypotheses that;

(1) for each group size, the mean assembly time decreases at a decelerating rate as a function of repetitions, and, after a certain number of repetitions, the mean assembly time would reach the steady-state level above zero (according to De Jong's (1957) Plateau learning curve model);

(2) for each repetition, the mean assembly time decreases at a decelerating rate as a function of group size; and,

(3) for each repetition, the mean productivity per worker decreases as a function of group size (Hypotheses 2 and 3 according to Steiner (1972)).

In addition to the hypotheses that were tested, the following main conclusions can be drawn from the present study;

- from a practical point of view, learning through repetition has a very significant effect on individual and group performance;

- larger groups resulted in increasing productivity loss in comparison to a one-worker group. Especially the workers in the four-worker groups experienced a loss of coordination, resulting from e.g. the limited physical space and insufficient number of subtasks for several workers. However, as the number of repetitions increased, the significance of group size for productivity decreased;

- though a minimum number of workers (one) resulted in the best productivity per worker, in the workers' opinions, a two-worker group is preferable, especially in connection with helping with larger parts. However, as the productivity varied significantly between different pairs, the selection of pair partners should be made carefully;

- with a relatively small number of repetitions (up to four), the workers from the larger groups (of three and four workers) mostly did not see too much harm in the group size and felt the size of their group was suitable. Thus, workers in larger groups seem to pay relatively little attention to declining efficiency with increased group size. Some of those from the three-worker groups, however, mentioned that after a certain number of repetitions a group of two workers would be preferable. In general, the opinion on an appropriate group size depends on the worker's own group size.

6.1 Further research

The results provided in this paper offer several topics for further research on the present case assembly. These topics are applicable to other assembly systems as well. For example:

- What are the underlying factors affecting fast learning and plateauing with different sizes of groups?

- What are the detailed sources of productivity losses with larger groups?

- What is the optimum size of a group in different assembly stages?

- What are the reasons for performance differences between worker pairs?

And perhaps most importantly, as assembly tasks in industry differ in terms of e.g. work contents and difficulty, and as real groups may operate quite differently, research should be moved ever closer to real production conditions.

References

Argote, L., S. L. Beckman, and D. Epple. 1990. "The Persistence and Transfer of Learning in Industrial Settings." *Management Science* 36 (2): 140–154.

Badiru, A. B. (Ed.). 2005. *Handbook of Industrial and Systems Engineering*. CRC Press. FL: Boca Raton.

Bukchin, Y., and Y. Cohen. 2013. "Minimising Throughput Loss in Assembly Lines due to Absenteeism and Turnover via Work-sharing." *International Journal of Production Research* 51 (20): 6140–6151. doi: 10.1080/00207543.2013.807374.

Buzacott, J. A. 2013. "Then and Now – 50 Years of Production Research." *International Journal of Production Research* 51 (23–24): 6756–6768. doi:10.1080/00207543.2013.778434.

Conway, R. W., and A. Schultz. 1958. "The Manufacturing Progress Function." Dept. of Industrial and Engineering Administration, Cornell University.

Dar-El, E. M. 2013. *Human Learning: From Learning Curves to Learning Organizations* Vol. 29. Springer Science & Business Media. New York. doi:10.1007/978-1-4757-3113-2.

Dar-El, E. M., K. Ayas, and I. Gilad. 1995. "A Dual-Phase Model for the Individual Learning Process in Industrial Tasks." *IIE Transactions* 27 (3): 265–271. doi:10.1080/07408179508936740.

De Bruecker, P., J. Van den Bergh, J. Beliën, and E. Demeulemeester. 2015. "Workforce Planning Incorporating Skills: State of the Art." *European Journal of Operational Research* 243 (1): 1–16. doi:10.1016/j.ejor.2014.10.038.

De Jong, J. R. 1957. "The Effects of Increasing Skill on Cycle Time and Its Consequences for Time Standards." *Ergonomics* 1 (1): 51–60. doi:10.1080/00140135708964571.

Fine, C. H. 1986. "Quality Improvement and Learning in Productive Systems." *Management Science* 32 (10): 1301–1315. doi:10.1287/mnsc.32.10.1301.

Frank, F., and L. R. Anderson. 1971. "Effects of Task and Group Size upon Group Productivity and Member Satisfaction." *Sociometry* 34 (1): 135–149. doi:10.2307/2786356.

Glock, C. H., and M. Y. Jaber. 2014. "A Group Learning Curve Model with and without Worker Turnover." *Journal of Modelling in Management* 9 (2): 179–199. doi:10.1108/JM2-05-2013-0018.

Hill, G. W. 1982. "Group Versus Individual Performance: Are N + 1 Heads Better than One?" *Psychological Bulletin* 91 (3): 517–539. doi:10.1037/0033-2909.91.3.517.

Ingram, P., and T. Simons. 2002. "The Transfer of Experience in Groups of Organizations: Implications for Performance and Competition." *Management Science* 48 (12): 1517–1533. doi:10.1287/mnsc.48.12.1517.437.

Jaber, M. Y. 2006. "Learning and forgetting models and their applications." *Handbook of Industrial and Systems Engineering* 30-1.

Jaber, M. Y. (Ed.). 2011. *Learning curves: Theory, models, and applications*. CRC Press. FL: Boca Raton.

Jaber, M. Y., and M. Bonney. 1996. "Production Breaks and the Learning Curve: The Forgetting Phenomenon." *Applied Mathematical Modelling* 20 (2): 162–169. doi:10.1016/0307-904X(95)00157-F.

Jaber, M. Y., Z. S. Givi, and W. P. Neumann. (2013). "Incorporating Human Fatigue and Recovery into the Learning–forgetting Process." *Applied Mathematical Modelling* 37 (12–13), 7287–7299. doi:10.1016/j.apm.2013.02.028.

Jaber, M. Y., and A. L. Guiffrida. 2004. "Learning Curves for Processes Generating Defects Requiring Reworks." *European Journal of Operational Research* 159 (3): 663–672. doi:10.1016/S0377-2217(03)00436-3.

Krishnaswamy, K. N., A. I. Sivakumar, and M. Mathirajan. 2009. *Management research methodology: integration of principles, methods and techniques*. Pearson Education. India: New Delhi.

Kuhrmann, M., and J. Münch. 2016. "When Teams Go Crazy: An Environment to Experience Group Dynamics in Software Project Management Courses." In *Proceedings of the 38th International Conference on Software Engineering Companion*. 412–421. ACM. doi:10.1145/2889160.2889194.

Li, G., and S. Rajagopalan. 1998. "A Learning Curve Model with Knowledge Depreciation." *European Journal of Operational Research* 105 (1): 143–154. doi:10.1016/S0377-2217(97)00033-7.

Lim, J. T. A., and E. R. Hoffmann. 2015. "Strategies in Performing a Manual Assembly Task." *International Journal of Industrial Ergonomics* 50 (1): 121–129. doi:10.1016/j.ergon.2015.06.005.

Littlepage, G. E. 1991. "Effects of Group Size and Task Characteristics on Group Performance: A Test of Steiner's Model." *Personality and Social Psychology Bulletin* 17 (4): 449–456. doi:10.1177/0146167291174014.

Nembhard, D. A., and N. Osothsilp. 2002. "Task Complexity Effects on Between-Individual Learning/Forgetting Variability." *International Journal of Industrial Ergonomics* 29 (5): 297–306. doi:10.1016/S0169-8141(01)00070-1.

Nembhard, D. A., and M. V. Uzumeri. 2000. "Experiential Learning and Forgetting for Manual and Cognitive Tasks." *International Journal of Industrial Ergonomics*, 25 (4): 315–326. doi:10.1016/S0169-8141(99)00021-9.

Niemi, E. 2009. "Worker Allocation in Make-to-Order Assembly Cells." *Robotics and Computer-Integrated Manufacturing* 25 (6): 932–936. doi:10.1016/j.rcim.2009.04.008.

Patel, H., M. Pettitt, and J. R. Wilson. 2012. "Factors of Collaborative Working: A Framework for a Collaboration Model." *Applied Ergonomics* 43 (1): 1–26. doi:10.1016/j.apergo.2011.04.009.

Peltokorpi, J., H. Tokola, and E. Niemi. 2015. "Worker Coordination Policies in Parallel Station Systems: Performance Models for a Set of Jobs and for Continuous Arrival of Jobs." *International Journal of Production Research* 53 (6): 1625–1641. doi:10.1080/00207543.2014.918290.

Qin, R., D. A. Nembhard, and W. L. Barnes II. 2015. "Workforce Flexibility in Operations Management." *Surveys in Operations Research and Management Science* 20 (1): 19–33. doi:10.1016/j.sorms.2015.04.001.

Ruiz-Torres, A. J., and F. Mahmoodi. 2007. "Impact of Worker and Shop Flexibility on Assembly Cells." *International Journal of Production Research* 45 (6): 1369–1388. doi:10.1080/00207540600690727.

Ryu, C., Y. J. Kim, A. Chaudhury, and H. R. Rao. 2005. "Knowledge Acquisition via Three Learning Processes in Enterprise Information Portals: Learning-by-investment, Learning-by-doing, and Learning-from-others." *MIS Quarterly* 29 (2): 245–278. doi:10.2307/25148679.

Saadat, M., M. C. Tan, M. Owliya, and G. Jules. 2013. "Challenges and Trends in the Allocation of the Workforce in Manufacturing Shop Floors." *International Journal of Production Research* 51 (4): 1024–1036. doi:10.1080/00207543.2012.662603.

Sando, T., M. Tory, and P. Irani. 2011. "Impact of Group Size on Spatial Structure Understanding Tasks." *IEEE Pacific Visualization*, 107–114. doi:10.1109/PACIFICVIS.2011.5742379.

Schilling, M. A., P. Vidal, R. E. Ployhart, and A. Marangoni. 2003. "Learning by Doing Something Else: Variation, Relatedness, and the Learning Curve." *Management Science* 49 (1): 39–56. doi:10.1287/mnsc.49.1.39.12750.

Sengupta, K., and F. R. Jacobs. 2004. "Impact of Work Teams: A Comparison Study of Assembly Cells and Assembly Line for a Variety of Operating Environments." *International Journal of Production Research* 42 (19): 4173–4193. doi:10.1080/00207540410001720421.

Sikström, S., and M. Y. Jaber, 2002. "The Power Integration Diffusion Model for Production Breaks." *Journal of Experimental Psychology: Applied* 8 (2): 118–126. doi:10.1037/1076-898X.8.2.118.

Slomp, J., and E. Molleman. 2002. "Cross-Training Policies and Team Performance." *International Journal of Production Research* 40 (5): 1193–1219. doi:0.1080/00207540110098823.

Staats, B. R., K. L. Milkman, and C. R. Fox. 2012. "The Team Scaling Fallacy: Underestimating the Declining Efficiency of Larger Teams." *Organizational Behavior and Human Decision Processes* 118 (2): 132–142. doi:10.1016/j.obhdp.2012.03.002.

Steiner, I. D. 1972. *Group Process and Productivity*. New York: Academic Press.

Thomas, E.J., and C. F. Fink. 1963. "Effects of Group Size." *Psychological Bulletin* 60 (4): 371–384. doi:10.1037/h0047169.

Uzumeri, M., and D. Nembhard. 1998. "A Population of Learners: A New Way to Measure Organizational Learning." *Journal of Operations Management* 16 (5): 515–528. doi:10.1016/S0272-6963(97)00017-X.

Wheelan, S. A. 2009. "Group Size, Group Development, and Group Productivity." *Small Group Research* 40 (2): 247–262. doi:10.1177/1046496408328703.

Whitney, D. E. 2004. *Mechanical Assemblies: Their Design, Manufacture, and Role in Product Development*. New York: Oxford University Press.

Wicker, A. W., S. L. Kirmeyer, L. Hanson, and D. Alexander. 1976. "Effects of Manning Levels on Subjective Experiences, Performance, and Verbal Interaction in Groups." *Organizational Behavior and Human Performance* 17 (2); 251–274. doi:10.1016/0030-5073(76)90066-0.

Wilson, J. M., P. S. Goodman, and M. A. Cronin. 2007. "Group learning." *Academy of Management Review* 32 (4): 1041–1059. doi:10.5465/AMR.2007.26585724.

Worchel, S., W. Wood, and J. A. Simpson. 1992. *Group Process and Productivity*. Newbury Park, CA: Sage Publications.

Wright, T. P. 1936. "Factors Affecting the Cost of Airplanes." *Journal of the Aeronautical Sciences* 3 (4): 122–128. doi:10.2514/8.155.

Yelle, L. E. 1979. "The Learning Curve: Historical Review and Comprehensive Survey." *Decision Sciences* 10 (2): 302–328. doi:10.1111/j.1540-5915.1979.tb00026.x

APPENDIX 1　　　　Assembly drawing

Item	Part/subassembly/small part	Qty	Tool
1	Frame	1	
2	Pipe subassembly 1	1	
3	Bolt, hexagonal head, M10x65, nut and 2 washers	2	2 x wrench 17
4	Module, 80x45x45	1	
5	Bolt, hexagonal head, M12x100, nut and 2 washers	1	2 x wrench 19
6	Pipe subassembly 2	1	wrench 36
7	Module, 70x70x70	1	
8	Screw, cylinder head, M5x75	2	hex key (yellow)
9	Module, 68x60x40	1	
10	Bolt, hexagonal head, M10x100, nut and 2 washers	2	2 x wrench 17
11	Pipe subassembly 3	1	wrenches 32 and 36
12	Hydraulic hose	1	wrench 24
13	Plate	1	
14	Screw, cylinder head, M5x60, nut and 2 washers	2	hex key (yellow) and wrench 8
15	Pipe subassembly 4	1	wrench 30
16	Hose subassembly, gray	1	
17	Screw, cylinder head, M4x25	2	hex key (green)
18	Pneumatic valve, 85x45x25	1	
19	Screw, cylinder head, M6x35, washer	2	hex key (red)
20	Hose, blue	1	
21	Pipe subassembly 5	1	wrench 30

APPENDIX 2   Principle of questionnaire

After each repetition:

Q1:   I tried to follow the processing pace according to the instructions (at a normal pace without hurrying).
1=completely disagree
2=somewhat disagree
3=neither agree nor disagree
4=somewhat agree
5=completely agree

Q2:   Two/three/four workers processing simultaneously caused inefficiency in group performance
1=not at all
2=rather little
3=neither much nor little
4=quite much
5=very much
If so, how did inefficiency occur in the performance?

After the second/third/fourth repetition:

Q3:   I learned from the previous repetition
1=not at all
2=rather little
3=neither much nor little
4=quite much
5=very much
How did learning occur in your own or the group's performance?

After the last repetition:

Q4:   From the performance point of view, the current group size was
1=too small
2=appropriate
3=too large
An appropriate group size would be __ workers.
Why would this group size be appropriate?

Q5:   The instructions were clear.
1=completely disagree
2=somewhat disagree
3=neither agree nor disagree
4=somewhat agree
5=completely agree
What was unclear?

Q6:   The exercise was comfortable.
1=completely disagree
2=somewhat disagree
3=neither agree nor disagree
4=somewhat agree
5=completely agree