Applied Ergonomics 58 (2017) 361–371

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Coordinating the interruption of assembly workers in manufacturing

Ari Kolbeinsson*, Peter Thorvald, Jessica Lindblom

University of Skövde, Box 408, 541 28, Skövde, Sweden

ARTICLE INFO

Article history:
Received 6 September 2015
Received in revised form
16 May 2016
Accepted 25 July 2016

Keywords:
Interruptions
Manual assembly
Manufacturing

ABSTRACT

This paper examines how interruptions from information and communications technology systems affect errors and the time to complete tasks for assembly workers. Interruptions have previously been examined in laboratory experiments and office environments, but not much work has been performed in other authentic environments. This paper contains the results of an experiment that was performed in a simulated manufacturing assembly environment, which tested the effects of interruptions on a manual assembly task. The experiment used existing interruption coordination methods as a basis, and the results showed a difference in the effect of interruptions and interruption coordination between cognitively complex laboratory tasks and manual assembly tasks in an authentic environment. Most notably, the negative effects of interruptions delivered without consideration were smaller in this experiment. Based on these findings, recommendations were developed for designing interruption systems for minimizing the costs (errors and time) imposed by interruptions during assembly tasks in manufacturing.

1. Introduction

The use of information and communications technology (ICT) has expanded greatly in recent years, having become near ubiquitous in people's personal lives and workplaces, including manufacturing facilities. This is exemplified by industrial frameworks and strategies such as Industrie 4.0 (e.g. Hermann et al., 2015), which focuses on the smart factories of the future, including both Internet of Things and Internet of Services. Smart factories inherently include extensive amounts of information flowing to and from users and one key component for successful utilisation of this information flow is the ability to know when and where information is needed as well as when and where additional information is not desirable. A key challenge in the smart manufacturing environments of today and tomorrow is to assess the current state of work and send various kinds of information to a user accordingly, as well as making a user aware of updated information that requires notifying the user that new information is available. This involves interrupting the user in some way and has been researched in many fields and domains for a long time, with the human factors advances in aviation being a prime example (e.g. Hawkins, 1993; Mirlacher et al., 2012). This has not been the case in assembly work where little research can be found on mitigating the effects of interruptions delivered by ICT systems on assembly workers. Manual assembly work on a production line often involves short and relatively simple tasks, with each workstation focusing on a minimum rational work element (Groover, 2010) as well as having an established time (takt time) in which each unit should be assembled (e.g. Womack and Jones, 2003).

Generally speaking, an interruption is anything that breaks into a user's current activity and demands a person's attention be shifted to another activity (Coraggio, 1990). It is of major importance to consider in what ways information interrupts and notifies workers in their assembly tasks in order to optimise work performance. Interruptions have been extensively investigated in several research areas (e.g. human-computer interaction (HCI), cognitive psychology, human factors) and differing domains (e.g. aviation, healthcare, office work) and potentially have large impact on work performance and output (e.g. Coraggio, 1990; Wickens, 1992; McFarlane, 2002; Iqbal and Horvitz, 2007; Iqbal and Bailey, 2008). However, it is notable that interruption research has not been applied to the manufacturing domain to any larger extent (but see Andreasson, 2014; Kolbeinsson et al., 2014), given its effect on work performance and work output. Interruptions can greatly affect workers' cognitive and mental load (e.g. Norman, 1993; Bannert, 2002) on both the primary task as well as the secondary task, depending on how and when the notification for the interruption is delivered (McFarlane and Latorella, 2002). Potential

* Corresponding author.
E-mail address: ari.kolbeinsson@his.se (A. Kolbeinsson).

http://dx.doi.org/10.1016/j.apergo.2016.07.015
0003-6870/

consequences of this include increased human errors, reductions in work output and a disregard for safety guidelines due to cognitive overload (e.g. Norman, 1993; Bannert, 2002).

The aim of this paper is to characterise the appropriate use of interruption coordination methods in manufacturing assembly and to highlight any differences from existing recommendations that have been developed using contrived and artificial tasks and environments. The main research question addressed regards what types of interruption coordination methods are suitable for use in manual assembly situations in manufacturing.

Much of previous interruption research in HCI has been carried out within controlled situations using contrived tasks and environments where the tasks are all manual, or tasks wherein the user sits in front of a stationary computer and both the primary as well as the secondary task happen on the same screen (e.g. McFarlane, 1999; Adamczyk and Bailey, 2004; Iqbal and Bailey, 2008; Grandhi and Jones, 2015). These studies also often use contrived, artificial tasks that are designed to set the difficulty, i.e. how much skill or effort is required to complete the tasks, of the primary task and interruption task high enough so that any increases in difficulty due to interruptions will result in errors being made (e.g. McFarlane, 1999; Adamczyk and Bailey, 2004; Grandhi and Jones, 2015). Iqbal and Bailey (2008) question whether results obtained using their contrived tasks can be applied directly to what they refer to as “authentic” tasks.

The contrived setups that have been described are useful for identifying fundamental interruption processes, but there is also a need for conducting applied research to complement the fundamentals in order to find a proper way to handle interruptions in other situations such as manufacturing and assembly where tasks are often simplified and optimised to avoid errors (e.g. Freivalds and Niebel, 2013; Brolin et al., 2012). Based on the research performed by McFarlane and Latorella (McFarlane, 1999, 2002; McFarlane and Latorella, 2002), where the fundamentals of interruption coordination methods were proposed and investigated, this work will elaborate on their findings and attempt to apply this on a more authentic scenario set in a simulated assembly context.

The intended contribution of this research is to identify and explore when, where and how to notify assembly workers of interruptions so as to minimise the negative consequences of interruptions. An over-arching goal for this work is more efficient management of assembly workers' cognitive load. Based on the obtained results, some recommendations for design of notification systems in manufacturing are provided that minimise increases in cognitive load due to interruptions.

2. Background

As research on interruptions has gone on for many decades within multiple fields (e.g. Coraggio, 1990; Wickens, 1992; Rubinstein et al., 2001; McFarlane, 2002; Iqbal and Horvitz, 2007; Iqbal and Bailey, 2008; Sykes, 2011; Warnock, McGee-Lennon and Brewster, 2011; Sanderson and Grundgeiger, 2015), this paper focuses on a subset of existing research, in particular on research that is relevant for work with mobile information devices used within the manufacturing domain.

Most manufacturing environments focus on the efficient mass production of products that requires supporting communications between managers, team leaders, and assembly workers (e.g. Bäckstrand, 2009). Timely dissemination of information can be vital for workers to complete their tasks, but unnecessary interruptions can have negative effects on performance and errors on the current task, and if the interruptions contain a new task to perform then interruptions at inappropriate times can also cause the interruption task to be completed with errors (e.g. Baron, 1986; Gillie and Broadbent, 1989; McFarlane, 1999; Zijlstra et al., 1999).

2.1. Interruptions and notifications

Coraggio (1990) defined an interruption as any external event that breaks into a user's current activity, the *primary task*, and demands the user's attention be shifted to another activity, the *interruption task*, or event (Coraggio, 1990). Interruptions are thus a very wide class of events, and can be anything from a random noise in the environment that causes the user to shift attention from the current task, to something that is specifically directed at a user for the purpose of diverting the user's attention through notifying that another task requires attention (Kolbeinsson et al., 2014). Interruptions may convey necessary information or superfluous information. Interruptions are referred to as *distractions* when they incur a measurable cost but do not result in a full switch from the primary task (Sanderson and Grundgeiger, 2015), which would include the example used of a random noise in the environment.

Interruptions can lead to more errors and longer time required to complete the primary task, as well as increasing stress and irritation due to increases in cognitive load (e.g. Wickens, 1992; McFarlane, 2002). These increases in cognitive load, stress, and irritation can also lead to more errors being made on both the interruption task and on the primary task (McFarlane, 1999). Directed interruptions commonly have the aim of supporting either the primary task or another task, and can thereby also be beneficial, bearing updated information so that the primary task can be completed correctly or supporting another task that must be completed. Interruptions can also be beneficial through raising cognitive load from a low state that may otherwise negatively affect performance through inducing boredom and inattention (Scerbo, 1998; Jackson et al., 2014).

The first known research on interruptions was published by Zeigarnik (1927), but research on interruptions was sparse after that until the rise of human factors research in the late 1970s (Spiekermann and Romanow, 2008). Increases in computing power and the development of more advanced computer systems then led to more complex office work and more requirements for ICT systems to interrupt workers (Speier, 1996). A consequence of this was a need for research on interruption management, which has mostly been conducted in lab environments (e.g. McFarlane, 1999, 2002; Altmann and Trafton, 2002; Iqbal and Bailey, 2008) as well as some observational studies performed to see what happens when interruptions occur in an authentic environment (e.g. Iqbal and Horvitz, 2007; Walter et al., 2015). Speier et al. (2003) found that interruptions have a larger negative effect on more complex tasks than on less complex tasks. Zijlstra et al. (1999) found that more complex interruptions result in more negative effects on the primary task, and Monk et al. (2002) showed that the difficulty of the primary task increases when the speed of the task is raised, with corresponding increases in negative effects of interruptions. The use of external cues also diminishes with increased complexity, i.e. tasks that may be more intricate and consist of a larger number of operations, or when shorter time is available to complete the task (Speier et al., 2003). This can affect the difficulty of the task and the quality of the work performed, with Speier et al. (2003) finding that participants performing tasks with a tight deadline make a trade-off in the quality of their work against performing the task in a timely fashion.

More research has been done on interruptions since, but as Brixey et al. (2007) as well as Walter et al. (2015) point out, this has mostly been carried out in laboratory experiments and may not be fully generalisable to authentic situations. Walter et al. (2015) have identified an interest in clarifying how interruptions affect occupational settings, and in particular stress the difficulty of quantifying the effects of interruptions in natural settings.

One major challenge in mitigating the costs of interruptions has been identified in the seminal works of McFarlane and Latorella as incorporating appropriate interruption coordination methods that include the costs of interruption into ICT systems (McFarlane, 1997, 1999, 2002; McFarlane and Latorella, 2002).

2.2. Interruption management

McFarlane (1999, 2002) and McFarlane and Latorella (2002) introduced a way of classifying different methods of interruptions with a taxonomy of interruptions that consists of four main ways of interrupting a user. They refer to these as the four methods of coordination, and are based on McFarlane 1999 experiment, with following work expanding upon that or further analysing results of McFarlane's experiments. The four interruption coordination methods explain differences in the points at which the user is notified of the new information held by the ICT system, the time at which the notification system interrupts the user and how directly the user is expected to respond to the interruption. These four coordination methods are (McFarlane, 1999, 2002; McFarlane and Latorella, 2002):

Immediate interruptions are delivered when the new information arrives and do not take into account current user activity.
Negotiated interruptions make the user aware that there is an interruption that will require some action, and may inform the user of the importance level of the interruption, but the user has control over when to pay attention to the notification and examine the new information itself in full detail.
Mediated interruptions use a mediator (a person or a ICT system) to gauge when it is appropriate to interrupt the user. There is a request that the user comes and performs an action, where this request is delivered via the mediator through the user's information device at a time deemed appropriate.
Scheduled interruptions are delivered at pre-scheduled times or intervals, which may take the importance of the interruption into consideration. Important interruptions might be shown every 5 min, while less important interruptions might be shown less frequently.

These four interruption coordination methods were identified through experiments done by McFarlane (1999, 2002) that consisted of a primary task involving subjects playing a computer game that involved catching people that jumped out of a burning building, and an interruption task using a modified Stroop task, which is used to set the difficulty level of the interruption task to avoid a floor effect for errors. Both tasks were performed on a single computer screen, with the interruption task completely replacing the primary task (MacFarlane, 1999, 2002). The primary (game) task required participants to correctly place a stretcher below a falling person, with each falling person bouncing so that the person must be caught three times by the stretcher. Each participant faced 59 such falling people in each of 24 trials. The interruption task, or matching task, showed a form with a colour (e.g. a blue triangle) and required the participant to select another element on the screen that matched either by shape or colour. Whether the match should be based on shape or colour varied. This allowed for fine grained control of the difficulty levels of both tasks as well as a clear switch between tasks, which made unambiguous data relatively easy to obtain as well as clearly showing the effects of interruptions at their most extreme through having only one task or the other visible at any given time. The design of the tasks was designed to be easy to measure and to isolate well the factors being measured; the tasks were not designed for authenticity or representativeness. Indeed, McFarlane states that "the difficulty of the tasks had to be contrived so that it was complex enough to attack participants' vulnerability to interruption, but simple enough not to cause participants to despair of performing well" (McFarlane, 1999, p.298).

Immediate interruptions automatically switched to the interruption task when the interruption was sent, occluding the primary task. Negotiated interruptions flashed the screen to show that the interruption task was waiting and then allowed participants to select a time they felt was appropriate for switching tasks. Mediated interruptions gauged how many jumpers were on-screen and waited for a suitable time with a low number of falling people, and scheduled interruptions were sent on a pre-arranged schedule of once every 25 s. Note that only one task or the other was visible at any given time regardless of which interruption coordination method was being used.

McFarlane (1999) findings showed all interruption coordination methods resulting in slower performance on the primary task than the control condition, with negotiated interruptions second fastest, mediated interruptions third fastest, immediate interruptions second slowest, and scheduled interruptions resulting in the slowest performance on the primary task. Errors on the interruption task were fewest on the control condition and most errors came from immediate interruptions. Mediated, negotiated, and scheduled interruptions did not show a significant difference from one another regarding errors on the interruption task, but showed significantly more errors than the control condition and significantly fewer errors than immediate interruptions. An important conclusion is that there is no single "correct" method of interruption (McFarlane and Latorella, 2002), as each interruption coordination method places different priorities on the primary and interruption tasks. McFarlane and Latorella (2002) considered multiple other factors such as the source of the interruption, channel of conveyance, individual characteristics of the user receiving the interruption, which all have bearing on which interruption coordination method to select.

Mediating interruptions was also shown by Arroyo and Selker (2011) to increase effectivity on a primary task as well as allowing a user to respond to more messages in the same time, thus providing a large gain in overall productivity. Knowing when to interrupt is clearly important, and ICT systems that know when to interrupt are important for safety, productivity, and human wellbeing. Iqbal and Bailey (2008) used McFarlane (1999, 2002) taxonomy for reference, and explored the granularity needed to find appropriate breakpoints for mediated interruptions during a document editing task, showing that sending interruption messages at breakpoints in activity decreased the costs of the interruptions. Although Iqbal and Bailey (2008) results were clear, they note that the tasks used were artificial in nature and might not apply to "authentic" tasks, and even if the findings apply to more authentic tasks then that is likely limited to similar tasks that involve document editing.

McFarlane (1999, 2002) and McFarlane and Latorella (2002) work has been widely used for reference in other work on interruptions, with the taxonomy of interruption coordination methods being used extensively (e.g. Iqbal and Bailey, 2008; Arroyo and Selker, 2011). Expanded taxonomies based on McFarlane (1999, 2002) also exist (e.g. Brixey et al., 2004; Rukab et al., 2004) but have not been as widely used. Minor concerns have been stated around McFarlane (1999, 2002) and McFarlane and Latorella (2002) taxonomy, arguing for its expansion to include additional factors, such as task switching and activity levels (e.g. Brixey et al., 2007). However, none of these concerns question the validity of McFarlane and Latorella's obtained results.

The same qualities that make McFarlane (1999, 2002) and McFarlane and Latorella (2002) work successful, i.e. the tight control of all variables, may create challenges in applying the results elsewhere (Rooksby, 2013). Many commonly performed tasks do not share the intrinsic difficulty level of a Stroop task, and additionally many common tasks are simplified as much as possible, with companies spending considerable resources on using production principles to simplify tasks, increase effectivity and decrease errors in assembly tasks in manufacturing (Bäckstrand, 2009). McFarlane (1999) stated that the difficulty of his tasks was designed to be just under what would make participants feel that the tasks were impossible. Likewise the task switching seen in McFarlane (1999) did not match many task switches in the real world, given that forcibly removing the task that is not in focus is not seen in most typical circumstances. For example, field observations made by Andreasson (2014) demonstrated how assembly workers that are interrupted select the time to respond to the interruption, and may even use their tools or protective gear as resumption points to assist in resuming their primary task. Baethge et al., (2014) addressed the need for more applied research on interruptions to complement contrived studies, thus broadening the scope of the field to also include the socio-technical environment. This suggests that some parts of McFarlane (1999, 2002) as well as McFarlane and Latorella (2002) have to be interpreted, re-examined, and validated for use in a more authentic environment.

This paper differentiates between tasks and environments designed to isolate and identify relevant factors, and tasks and environments that are designed to elicit responses similar to those shown by people when performing tasks in their day-to-day lives, whether that is in their work or in their personal time. The prior is referred to here as a *contrived* task or environment, and the latter as an *authentic*, or *more authentic* task or environment. The use of these terms is inspired by Iqbal and Bailey (2008) note that the behaviours observed in experiments using contrived tasks may not be generalizable to more authentic tasks.

2.3. Summary and hypotheses

The controlled nature of McFarlane (1999) experiment suggest using authentic tasks and environments, as opposed to the contrived tasks used by McFarlane, to support the creation of systems for minimising the negative effects of interruptions on users. As the focus in this work lies on assembly workers in manufacturing then this suggests the use of either an experiment or a quasi-experiment in a manufacturing facility. The complexity, cost, safety factors, and quality control factors make this unfeasible, leading to the use of a simulated factory environment where appropriate task and domain factors are simulated to create the level of engagement and realism required (Drews and Bakdash, 2013; Sanderson and Grundgeiger, 2015). The study investigates whether the interruption coordination methods as described by McFarlane (1999) as well as McFarlane and Latorella (2002) result in similar findings when performed in a simulated manufacturing environment using real assembly tasks. Differences in results from McFarlane (1999) are predicted to be that the authentic assembly tasks will not generate as many errors on either the primary or the interruption task, that errors will not be increased as much as shown by McFarlane (1999), and that immediate interruptions are predicted not to cause nearly as large negative effects as was observed by McFarlane (1999). Immediate interruptions are expected to generate less costs than predicted by McFarlane (1999) due to the task switching control being in the hands of the participants, as opposed to McFarlane's use of task occlusion, where the primary task was removed from the screen and replaced with the interruption task which leaves task switching control firmly in the hands of the researcher. Other interruption coordination methods retain their primary characteristics when applied to the simulated manufacturing assembly tasks and environment.

The hypotheses are as follows:

- The interrelations observed between mediated, negotiated, and scheduled interruptions will correspond to those observed by McFarlane (1999).
- Immediate interruptions will result in less relative costs than observed by McFarlane (1999).

The anticipated differences from McFarlane (1999) work are all expected due to the nature of the more authentic tasks and environment, including the lack of task occlusion, as compared to the artificial tasks used by McFarlane (1999), task switching control residing more with the participants, as well as the difference between the cognitive complexity of the artificial lab tasks used and more common tasks which may be optimised for lower cognitive complexity.

3. Method

To test the hypotheses, the original 1999 experiment by McFarlane was adapted through the use of a more authentic scenario and environment than that used by McFarlane (1999). The tasks and context were both designed specifically for this experiment but McFarlane (1999) basic approach, aim, and results were used for reference. This means that the experiment in this paper was not a replication, but rather an adaptation that used McFarlane (1999) and McFarlane and Latorella (2002) as a starting point. The purpose here was thus not to replicate McFarlane (1999), but rather to expand upon McFarlane (1999) detailed findings. The experiment involved participants performing a repetitive manual assembly task involving mounting the two front wheels onto pedal cars, receiving an interruption on a mobile device which involved having to go to a stationary information system to read error messages, interpret those, and sending the correct response to the correct person as indicated by the message on the mobile device.

Performing the experiment in an actual manufacturing assembly facility would have been ideal, but was found to be impractical due to the level of control required to manipulate the variables. This led to the examination of relevant alternatives, with the chosen approach using a simulation of the task and context being studied.

3.1. Research design

In this paper, the concept of simulation is aligned with the complementing views proposed by Rooksby (2013) as well as Sanderson and Grundgeiger (2015), focusing on recreating a context and task to an appropriate degree of *realism* (accuracy of the physical and psychosocial environment) and *fidelity* (the visual and technical accuracy of the simulation). Other forms of simulation exist, but will not be further detailed here. Rooksby (2013) argues the use of simulations allows laboratory studies to be performed in such a way as to elicit similar behaviour as would be observed in a real task and context through designing laboratory tasks and environment to take into account human practice. Drews and Bakdash (2013) further explain that simulations for the purpose of research should support a high level of engagement and be easy to manipulate, pointing out that this is more important than focusing on ultimate fidelity through perfect recreation of the target domain.

An appropriate level of fidelity was achieved in this study through designing the primary task to closely resemble tasks observed in manufacturing assembly situations, and through making the UI for the interruption task resemble flawed communications systems which have been observed by the authors in several industrial/manufacturing applications. The background story, the configuration of the locale, as well as visual and auditory aspects of the context were created to support the validity of the research through increasing the immersiveness of the context, as suggested by Tolmie and Crabtree (2008). This is further supported by Sanderson and Grundgeiger (2015) who refer to immersiveness as the *representativeness* of the task and context. Feedback from a small pilot study supported that the tasks and context achieved the desired effect.

3.2. Experimental design

As the experiment examined effects seen in McFarlane and Latorella (2002), the independent variable was known from the beginning to be the *interruption coordination method*, and the dependent variables for the experiment also had to be comparable to McFarlane and Latorella (2002) yet support the new scenario. The independent variable consisted of six levels that created the experiment conditions and included the four interruption coordination method treatments as well as two control conditions. The conditions were:

- Immediate interruption (Imm.)
- Negotiated interruption (Neg.)
- Mediated interruption (Med.)
- Scheduled interruption (Sched.), every third minute
- No interruption, assembly only (Assem. only)
- Only an interruption, no assembly (Int. only)

The dependent variables measure the cost of the interruption, and were:

- Assembly time (Assem. time): time to complete a single car assembly (seconds)
- Assembly errors (Assem. errors): number of errors made on the assembly task
- Message time (Msg. time): time to respond to the interruption task message (seconds)
- Message errors (Msg. errors): number of errors made on the interruption message task

A within subjects, repeated measures experimental setup as used, with participants assigned to one of six groups and balanced against training effects through the use of a Balanced Latin Square setup identical to that used by McFarlane (1999).

The experiment was performed in an environment that is used for teaching production principles to industrial assembly workers and which was specifically modified for this experiment. The physical environment can be seen in Fig. 1. A slightly manipulated audio recording from a real manufacturing facility was played at high volume during the experiment to create a realistic aural environment. Hearing protectors and brightly coloured safety vests were worn to make the simulation as immersive and authentic as possible.

3.2.1. The primary task

The primary task (car assembly) involved mounting two front wheels onto a pedal car, with the small extra complication that each car required a specific setup of tire hardness (three levels) and each wheel had specific markings denoting hardness. The steps required to complete the task were: selecting and mounting the correct wheels onto the axle, threading a bolt and washer on the axle, fastening using a cordless drill. Each pedal car had a note on the back stating which hardness was needed for each front wheel. The note had the tire hardness written above and below one another, to increase the difficulty of translating to left and right through visual mapping. A confederate in another part of the facility was tasked with disassembly of the pedal cars and recording assembly errors such as wrong wheel hardness or errors on bolt/washer assembly or loose bolts.

Fig. 1. The environment for the experiment seen from the same point as the video recordings. The assembly station and touch screen can be seen, as well as the cars awaiting assembly. The computer system and assembly station can be seen on the right side of the picture, and the disassembly station can be seen on the left side.

The design of the primary task was based on tasks observed in manual assembly work, and verified as being representative of the domain by researchers with experience from manufacturing assembly. It should also be noted that as soon as one car had been assembled, another car arrived at the assembly station immediately and more cars awaited assembly in line, see Fig. 1. The primary task had a similar time for assembly (takt time), complexity, and requirements for manual coordination as assembly operations observed in a white goods manufacturing facility. The short time allowed for the primary task affected the difficulty of the task (Monk et al., 2002).

3.2.2. The interruption task

The interruption task involved a message sent to a mobile device carried by the participants, which asked for specific error codes to be sent to specific recipients, based on a cover story introduced at the beginning of the experiment. The message was sent using a pre-production version of a messaging system in development for commercial applications. The system contained a random timer, which prompted the researcher to send the message and create a timing event in the software. When responding to a message, participants were required to walk to a stationary terminal (situated approximately 4 m away from the assembly station), read error information from the message on the mobile device, find the correct error code for that error on the stationary screen, select the matching symbols on the stationary terminal, and send to the correct recipient. The error codes as displayed involved a sequence of four basic shapes written out as words and inputting the symbols into the system then required participants to press actual symbols matching the text. This was originally inspired by the Stroop test (in a similar manner to McFarlane, 1999), but was redesigned to more accurately represent tasks faced by manufacturing assembly workers while remaining cognitively demanding. This task can be seen in Fig. 2. The discrepancy between text and symbols was used to hinder participants from using visual pattern matching to input codes by creating a requirement for participants to perform an

Fig. 2. The touch screen interruption task system. The instructions must be found in the text and then input through the symbols and selections on the bottom half of the screen.

interpretation between words and shapes and increase the memory requirements of the task (Floel et al., 2004). Messages were sent in a pre-defined sequence, and participants' responses were recorded by the system so that timing and correctness of messages could be verified. The interface for sending the error codes was designed based on the interface problems observed by the authors in multiple manufacturing facilities in several countries. The design of the user interface was performed by a user interface specialist, and broke against best practices for the reasons outlined above. The interruption messages were slightly abstracted, but were designed to be thematically related to the frame of the study, which Sanderson and Grundgeiger (2015) suggest as an appropriate approach for use in simulations to maintain the appropriate level of fidelity in the simulation.

Careful consideration was taken to structure both the primary task and the interruption task in a way similar to what is seen in manufacturing facilities, and control over task switching was left in participants' hands. This was done to ensure that the tasks were as authentic as possible within the simulation to maintain representativeness.

3.2.3. Equipment

The facility used had multiple assembly and disassembly stations, each with the correct manual tools or power tools as required. A single assembly station was used by the participant, and a single disassembly station was operated by a confederate. Custom computer programs were used to record data from the primary task, the interruption task, to send interruption messages, and to customise the messages that went with the interruptions.

The messaging system controlled the messages sent to the mobile device in the participants' pocket, including intensity and persistence of both audio and vibration. Immediate interruptions used a loud continuous sound and vibration which had to be dismissed by the user while all other interruption coordination methods used a single notification in both sound and vibration.

A touch screen controller (see Fig. 2) was used for the input of participants' responses to the interruption task. This was used to simplify the physical part of the data entry by giving direct input so as to eliminate any confounding factors arising from a user interface (UI) using a relative and translational interface such as a mouse (Cooper et al., 2014).

3.3. Participants

Participants were recruited through advertisements in the university and in the surrounding town. 25 people participated. Close proximity to major manufacturing facilities meant that nine participants had work experience within manufacturing, ranging from a few months of experience to 26 years of experience. Participants' age ranged from 18 to 53 years, with a majority, 17 participants, between the age of 25 and 40. The participant pool was deemed as being fairly typical of new hires in local manufacturing, where many locals work for some periods of time before pursuing further education or specialisation. Seven participants were female and 17 male. Gender comparison was not the focus in this experiment, and was deemed impractical due to small sample size. Gender difference was further deemed unlikely to affect results as McFarlane (1999, 2002) showed no significant difference between genders. Risks were explained, safety equipment used, and consent obtained both for participation and for using video recordings. The experiment ended with a debriefing that included a post-test questionnaire for recording participants' subjective experience from the experiment. Each participant received a practice period until the participant performed the primary and interruption tasks without procedural errors at least three times in a row.

As the experiment examines a task involving unskilled labour using only basic tools, the training period was deemed as being sufficient to maintain a fidelity high enough to support the aims of the study. This was controlled through prior testing of the tasks by the researchers and verification through the use of a pre-study that used unskilled participants. The behaviour and performance of participants with assembly experience versus those without was also observed throughout the experiment. No behaviour differences or performance deviations were observed based on experience levels.

3.4. Procedure

Each interruption coordination method was tested separately, having each participant perform the assembly task until six interruption measurements were verified as having been successfully recorded, which involved a minimum of six assembled cars per participant for each condition. This took approximately 45 min for each participant, during which time participants could not leave the assembly station as a constant flow of assembly tasks simulated authentic manufacturing environments. The assembly only (assem. only) control treatment involved a participant performing the primary task (assembling cars) without interruption, and the interruption only (int. only) control treatment involved a participant standing without the primary task (assembly) in a designated area that matched the distance that participants had to walk from the primary task to the touchscreen system, acting upon a message of the same type as in other treatments, and walking back to the designated area.

Immediate interruptions were sent based on a signal from the random timer in the experimental software. Immediate interruptions were loud and persistent, that is, continued until participants pressed a button on the phone. Participants were instructed to respond immediately to these interruptions.

Mediated interruptions were sent at one of two pre-defined breakpoints: after mounting one wheel but before starting on mounting the second wheel (between wheels), or after the assembly of one pedal car was complete but before starting the assembly of the next pedal car (between cars). These breakpoints were selected based on Iqbal and Bailey (2008) who showed that interruptions between tasks or sub-tasks are less deleterious than random interruptions, and that coarser granularity of breakpoints results in better performance. The random timer was used to select which breakpoint was used.

Negotiated interruptions were sent at random times using the random timer, but participants were instructed to respond at a time of their own choosing.

Scheduled interruptions involved messages being sent at random times using the random timer, but the notification message was sent every 3 min. A clock was also mounted in a highly visible position to the participants with markings showing the minutes at which interruptions would arrive.

4. Results

The data for each dependent variable was analysed using ANOVA for finding significant effects, effect sizes, and power. Visualisations were also extensively used to understand the data, as per Cohen (1994) and Tukey (1962) recommendations for avoiding common mistakes in interpreting results. Cohen (1994) advice to carefully examine effect sizes, confidence intervals, and observed power to minimise the risk of type II errors was also followed.

Missing values for assembly time in the scheduled condition were found to be 10.5%. Other conditions had less than 1.5% values missing. This was due to an issue with how data was sampled during scheduled interruptions, which did not take into account the extra time required around the scheduling times, which were every third minute. This meant that fewer cars were assembled during scheduled interruptions. Imputation was used to fill in the missing values. Imputation maintains the variance as well as the means, thus affecting the results less than using the means in the case of many missing values (Donders et al., 2006). The use of imputation to fill in missing values lowers the robustness of the results for assembly time in scheduled interruptions, but increases the number of valid measurements that could be used from all other conditions. One participant was removed from the data as the participant made systematic errors accounting for over one third of all errors in the entire experiment, as errors were otherwise infrequent. All effect sizes (η_p^2) shown in Table 1 take into account the appropriate correction for sphericity.

4.1. Effects of interruptions

Overall effects of interruptions (see Table 1) were observed, and were examined for both the primary task and interruption task as well as for each interruption coordination method. Pairwise comparisons were used for identifying effects between individual interruption coordination methods.

4.1.1. Primary task (Assem. time and Assem. errors)

Assembly time (Fig. 3a) showed a significant main effect between the control condition (assem. only) and others in the pairwise comparison. All interruption coordination methods were seen to be significantly slower than the control, but no significant difference was thus observed between interruption coordination methods.

No significant effects were observed for *assembly errors*. Fig. 3b shows the means and the error bars show the upper and lower bounds of the 95% confidence interval. The means may seem to suggest an effect and show exactly what was expected, i.e. that the assembly only (assem. only) control treatment resulted in the fewest assembly errors, and scheduled interruptions were similar, but the high variance and low total number of errors made means that these differences are not statistically significant. Assembly errors thus show no effect based on interruption coordination method.

4.1.2. Interruption task (Msg. time and Msg. errors)

Significance was observed on the *message time* (Fig. 3c) variable between all interruption coordination methods and the control except between immediate interruptions and the interruption only control treatment (int. only). Immediate interruptions were thus as fast as the interruption only control condition, confirming that participants did respond immediately, as instructed.

Significance on the *message error* (Fig. 3d) variable was only observed between the immediate and negotiated treatments ($p = 0{,}037$, $\alpha = 0{,}05$). This was only seen after performing imputations, but was seen reliably in imputed data. The data also suggests that an effect may exist between immediate and mediated, as well as possibly between immediate and scheduled interruptions, but the high variance and low number of total errors mean that no significance was observed and no claims can be made based on this.

4.2. Results of hypotheses

The first hypothesis was:

- The interrelations observed between mediated, negotiated, and scheduled interruptions will correspond to those observed by McFarlane (1999).

As has been demonstrated by the data, mediated, negotiated, and scheduled interruptions do not differ significantly from one another on assembly time, assembly errors, or message errors. However, there is a significant difference between the three when it comes to message time. This does not correspond with McFarlane (1999, 2002) findings. Thus the hypothesis is rejected.

The second hypothesis was:

- Immediate interruptions will result in less relative costs than observed by McFarlane (1999).

As has been demonstrated by the data, immediate interruptions do not differ from the other three interruption coordination methods on assembly time or assembly errors. Immediate interruptions differ significantly from all other interruption coordination on message time, as well as showing significantly fewer message errors than negotiated interruptions. Thus the hypothesis is supported.

Table 1
Statistics for the main effects of the four dependent variables.

Main effect significance table	P value	Partial eta squared (η_p^2)	F	Fcrit	Obs. power
Assembly time	0,000	0272	42,929	2600	1000
Message time	0,000	0688	254,095	2640	1000
Assembly errors	0,101	0017	1951	2480	0,588
Message errors	0,069	0019	2201	2480	0,643

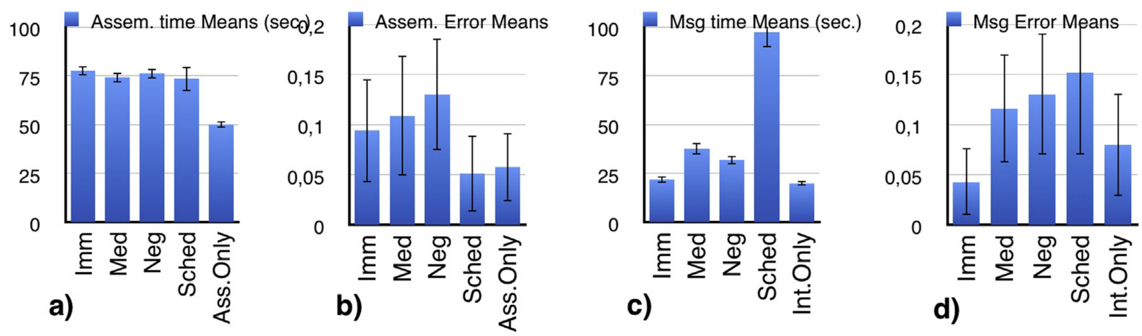

Fig. 3. Means and error bars showing the 95% confidence intervals for the four dependent variables.

4.3. Further findings

In addition to the results of the hypotheses above, this section presents further findings relevant to the overall research question.

Firstly, immediate interruptions were observed to not affect performance on the primary task to any similar degree as observed by McFarlane (1999) as well as McFarlane and Latorella (2002), and resulted in fast and accurate responses to the interruption. Contrary to expectations, immediate interruptions were seen to perform extremely well, with no obvious effect on primary task performance, no significant increase in errors on either primary or interruption tasks, and the fastest response times to the interruption task. This is a stronger effect than expected, but still along the lines of the second hypothesis. However, immediate interruptions were reported by participants as being stressful, and a participant suggested these might become difficult to handle if used for a longer time due to their stressful nature.

Secondly, mediated interruptions were observed creating unexpected issues. The mediation was controlled by using preselected points in the assembly that were expected to minimise the cost of interruption, using a random timer to select which preselected point to interrupt at, and sending a message at that point. The pre-selected points were (1) after finishing the assembly of one wheel but before starting work on the second wheel, or (2) after both wheels had been mounted but before starting work on the next pedal car. These two points are referred to as *between wheels* and *between cars*. These points were selected as being between tasks, which should, according to Adamczyk and Bailey (2004) as well as Iqbal and Bailey (2008) result in participants being disturbed less and performing better than when being interrupted during a task to make resuming the primary task easier. It was observed that when interrupted between wheels, a large number of participants missed the interruption notification entirely, while messages between cars did not suffer this problem. This observation suggests that mediated interruptions are heavily dependent on selection of appropriate interruption points and requires further examination. These missed notifications were counted as errors, and thus increase the message error rate measured on mediated interruptions. This results in mediated interruptions showing a worse performance in both time and errors than if interruptions had been sent at optimal breakpoints.

Furthermore, an element of negotiation was observed regardless of which interruption coordination method was used. The use of immediate interruptions resulted in the fastest response to the interruption, but participants still responded after finishing their current operation, which is a form of negotiation. Mediated interruptions attempted to interrupt at convenient times and when that was successful, i.e. between cars, participants responded quite fast but still often looked at the car to see that it was finished and pushed the car away to the next station to prepare for the next car before responding to the message. When faced with scheduled interruptions, participants would check the clock to see how long they had, and plan their assembly in accordance with that, to a degree. Even so, response to the interruption was delayed so that participants could finish their current task. Moreover, negotiated interruptions were mentioned as preferred by participants in the debriefing.

Finally, scheduled interruptions resulted in the longest response time for the interruption task, but less expectedly, scheduled interruptions were not rated positively by participants in the debriefing. This was, according to participants' comments, due to the complexity of having to deal with multiple messages at the scheduled times.

5. Discussion and conclusions

The aim of this paper was to characterise the appropriate use of interruption coordination methods in manufacturing assembly and to highlight any differences from existing recommendations that have been developed using contrived and artificial tasks and environments.

The main research question addressed regards what types of interruption coordination methods are suitable for use in specific assembly situations in manufacturing. In order to do that, we adapted McFarlane (1999) research to the manufacturing domain. Findings indicate that the more authentic setting of the experiment has an effect on the outcome, giving different results to those obtained by McFarlane (1999, 2002).

5.1. Discussion

The obtained results mostly support McFarlane (1999) original claims, with the exception of the lower negative impact of immediate interruptions than observed by McFarlane (1999, 2002), as hypothesised. Based on the obtained results, the first hypothesis is rejected, as the difference in time to complete the assembly task showed no statistically significant difference between interruption coordination methods. A possible explanation for this finding is that physical nature of the tasks used allow for the use of spatial orientation and other salient cues to resume the assembly task, i.e. that participants could remember and see *where* to continue, thereby creating an easy resumption point for *what* to do next.

The second hypothesis is supported by the obtained results, but the unexpected findings encountered require further examination and discussion. Although the second hypothesis predicted that immediate interruptions would not cause as much disruption as seen by McFarlane (1999, 2002) the results showed immediate interruption not causing a measurable cost in time or errors *at all* when compared to other interruption methods on the primary (assembly) task. This supports the reasoning behind the second hypothesis, although the effect is stronger than expected. The hypothesis was formed on the grounds that more authentic tasks would require less complex mental work than the lab based tasks used by McFarlane (1999, 2002), as well as immediate task switching in more authentic tasks working in a different manner to that employed by McFarlane (1999, 2002). This is supported by the stronger than expected results.

A second factor that may have contributed to immediate interruptions performing well is the participant controlling the task switching during authentic tasks as compared to lab-based tasks, where the experimenter controls the task switching. Task switching in more authentic contexts commonly involves becoming aware that it is necessary to change tasks, changing tasks, completing the secondary interruption task, and switching back to the primary task to complete it. This differs in a central way to McFarlane (1999) approach, which essentially instantly removed the primary task in the immediate condition, i.e. used task occlusion. As previously explained, this is not a problem with McFarlane (1999) approach, but should rather be viewed as a feature that allowed McFarlane (1999) to better isolate the independent variables of the experiment. However, the same feature may not translate to many more authentic tasks, which is a reason for the more authentic tasks and environments used here being more appropriate for supporting the design of interruption coordination systems for the manufacturing assembly domain.

The difference between McFarlane (1999, 2002) lab experiment and the approach used here is further highlighted by an effect that we have named the *negotiation element*. Most authentic tasks allow a user some time to finish or suspend a task before switching to another. Negotiated interruptions are an example where this is the overriding principle, and the observations made in this experiment is that most authentic tasks allow for some amount of negotiation. The amount of negotiation can vary, but tasks where people instantly let go of the task (and tools) they are currently performing are rare and usually have to do with emergencies rather than repeated tasks. This effect should be considered when creating interruption systems for real tasks. The negotiation element requires more study, as this may affect resumption and errors in authentic tasks.

Unexpected findings observed during mediated interruptions were interesting, as prior research conducted by Adamczyk and Bailey (2004) as well as Iqbal and Bailey (2008) suggests that breakpoints at the end of a main task minimise task-switching costs, and that breakpoints between sub-tasks reduce task-switching costs slightly less. The problems observed saw participants missing interruption messages sent “between wheels”. This suggests that something more is going on in the tasks tested, as participants frequently missed interruption messages sent between sub-tasks (“between wheels”). One possible explanation for this is that the physical task switching and mental task switching are performed asynchronously and that the breakpoint for participants' mental task switching may have already passed at the physical task breakpoint that was selected. This requires further study.

The experiment and its results clearly show the differences between testing for an effect using artificial tasks that are designed to elicit a certain response compared to using more authentic tasks and environments that support human practice. As such, the experiment supports Rooksby (2013) view that there is a need for adding the use of simulations to the research toolbox, as traditional experiments may not accurately predict performance in many cases. Furthermore, Rooksby (2013) support for the use of simulation is supported, so long as the tasks and environments are designed with an understanding of the required fidelity and realism (Smallman & St. John, 2005; Sandersson and Grundgeiger, 2015).

McFarlane's notes on the tasks used in his experiment vividly highlight the difference between the tasks used in his 1999 experiment and the simplified tasks more commonly faced in assembly work. This supports our claim that neither task was representative of everyday tasks that people do, and thereby not representative of most work activities such as assembly tasks in manufacturing. McFarlane increased the task complexity in his experiment by increasing the cognitive requirements of the tasks involved through setting the speed of the primary task (catching people falling out of windows in a computer game), and through the use of a task known for its cognitive complexity (a modified Stroop task) for the interruption task (McFarlane, 1999).

An important lesson learned is that building any kind of interruption coordination mechanism into an ICT system is optimally based on research that uses a similar or identical task to that which should be supported by the system, and that the *negotiation element* that is present in most types of task switching should always be taken into account. This lack of a negotiation element is offered here as a tentative explanation for the good performance observed when using immediate interruptions.

Participants reported the experiment as being immersive, and two participants with prior experience from assembly work in manufacturing made unsolicited comments stating that the experimental setup felt like being back at work. Observations supported this, as participants were seen to do their best and responded as if the tasks were real, suggesting a high level of engagement, which was further observed when participants noticed themselves making errors. Unsolicited comments made by participants suggest a high level of representativeness of the tasks and environments, and further support assumptions made by the authors at the outset. The authors noted that the environment felt representative of a manufacturing assembly facility from the standpoint of their role as observers and confederates, with the context feeling immersive for the authors and participants, and the soundscape in particular adding to the illusion of a fully functioning manufacturing assembly facility.

This supports the chosen approach of using a simulated assembly line with artificial tasks that emulate tasks seen in manufacturing. The primary task might be seen as being simple, but the task was designed to represent common assembly tasks in manufacturing. Assembly tasks in manufacturing are generally kept as simple as possible, but still differ considerably based on what is being manufactured. Assembly of car engine parts may thus have a higher level of complexity in its minimum rational work element, than the assembly of a washing machine vibration damper. Experimenting with different levels of complexity of primary tasks may therefore give different results. According to Sanderson and Grundgeiger (2015) potential generalizability refers to two things, the potential for depth of insight, and breadth of application of conclusions. The level of engagement, and the apparent representativeness of the simulation used in the study suggest that the potential for depth of insight is high, which allows us to create guidelines for the design of interruption coordination systems for use within manufacturing assembly situations. The limitations on potential generalizability, specifically in terms of breadth of application, are that we cannot claim that the results of this study can be applied to other domains or other forms of tasks.

The low number of participants, 24, is partly offset by the repeated measures design used, but more participants would make the results more robust. The problems seen in sampling data for the scheduled interruption condition also affected the robustness of that condition, but these problems do not affect other conditions. Any replication of this experiment, or further work, should therefore approach data sampling for the scheduled condition in a slightly different manner, ensuring that the appropriate number of samples is recorded each time instead of allowing time limits to impinge on the data sampling.

5.2. Conclusions and recommendations

The main contributions of this work are firstly the finding that immediate interruptions result in much less negative effects when these interruption coordination methods are used in more authentic assembly tasks than was seen in prior research that relied on more artificial tasks. The stressful effects of immediate interruptions that have been observed in prior research were also reported here, but did not translate to participants' performance. Secondly, mediated interruptions were seen to be vulnerable to the selection of breakpoints, with notifications being missed or ignored if sent at an inappropriate time. Thirdly, we have introduced the concept of a negotiation element, which is arguably present in authentic situations but may be absent in artificially contrived situations. Limited research exists using authentic tasks, which means that finding research that matches the intended context of use for an intended ICT system is rare.

Based on these contributions we have developed a set of recommendations for the design of interruption systems for supporting assembly work in manufacturing, which in the long run may reduce assembly workers' cognitive load.

- Immediate interruptions are an acceptable choice for manual tasks and when interruptions are infrequent, as well as for when interruptions should be prioritised over performance on the primary task. Only scheduled interruptions resulted in less negative effects on the primary task.
- Mediated interruptions are only useful if interruption points have been identified and verified as being appropriate for the task at hand. Mediated interruptions should be selected when interruptions are more frequent and when the interruption task priority is lower than the primary task priority, i.e. when the interruption task does not require immediate action or response.
- Negotiated interruptions are appropriate when interruption task priority is low, but primary task priority is medium (not critical). Reminders should be used to avoid workers forgetting to respond.
- Scheduled interruptions should be used for low priority messages or where the time to respond to the message is not a factor and where primary task performance is prioritised high or critical. Scheduled interruptions are, for example, appropriate for sending messages shortly before or after a scheduled worker break.

These recommendations are a first step in creating an integrated interruption management framework for minimising errors and assembly time increases through managing cognitive load in manufacturing assembly situations.

5.3. Further research

Follow up studies have been planned based on the unexpected observations made on mediated interruptions. Some further analysis of the obtained results has been conducted from an embodied cognitive science perspective (Kolbeinsson and Lindblom, 2015), but this is only a first step and requires further work.

Furthermore, a need has been identified to introduce other explanatory concepts and frameworks that incorporate a larger unit of analysis, rather than separating the primary and secondary tasks from the actual work situation environment (workstation design and work environment), including trigger analysis (e.g. Dix et al., 1998; Kirsh, 2001), distributed cognition (e.g. Hutchins, 1995; Hollan et al., 2000) and activity theory (e.g. Kaptelinin and Nardi, 2006) in order to consider the broader socio-technical context, in which interruptions and assembly workers are situated in manufacturing. There are huge costs associated with neglecting cognitive and user perspectives within manufacturing, but on the other hand, there is a vast potential to improve both the workers' cognitive load and an increased production outcome simultaneously if interruptions are handled properly. We propose that an integrated socio-technical approach for handling and coordinating interruptions of assembly workers may provide promising and necessary, but not sufficient, steps towards realizing the smart factories of the future.

Acknowledgments

The authors wish to thank the participants in the study as well as the Sense&React Consortium for their input and support. Special thanks to Per-Anders Johansson and Magnus Nordlund at Ascom for assistance with equipment and software. The Sense&React project is funded by the EU grant FP7-314350.

References

Adamczyk, P.D., Bailey, B.P., 2004. If not now, when?: the effects of interruption at different moments within task execution. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 271–278.

Andreasson, R., 2014. Interruptions in Manufacturing from a Distributed Cognition Perspective (Master's thesis). University of Skövde, Skövde, Sweden.

Altmann, E.M., Trafton, J.G., 2002. Memory for goals: an activation-based model. Cogn. Sci. 26 (1), 39–83.

Arroyo, E., Selker, T., 2011. Attention and intention goals can mediate disruption in human-computer interaction. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (Eds.), Human-computer Interaction – INTERACT 2011, 13th IFIP TC 13 International Conference (Lisbon, Portugal, September 5-9, 2011): Proceedings, Part II. Springer, Heidelberg, pp. 454–470.

Baethge, A., Rigotti, T., Roe, R.A., 2014. Just more of the same, or different? an integrative theoretical framework for the study of cumulative interruptions at work. Eur. J. Work Organ. Psychol. 24 (2), 308–323.

Bannert, M., 2002. Managing cognitive load – recent trends in cognitive load theory. Learn. Instr. 12, 139–146.

Baron, R.S., 1986. Distraction-conflict theory: progress and problems. Adv. Exp. Soc. Psychol. 19, 1–39.

Brixey, J.J., Walji, M., Zhang, J., Johnson, T.R., Turley, J.P., 2004. Proposing a taxonomy and model of interruption. In: Enterprise Networking and Computing in Healthcare Industry, 2004. HEALTHCOM 2004. Proceedings. 6th International Workshop on. IEEE, pp. 184–188.

Brixey, J.J., Robinson, D.J., Johnson, C.W., Johnson, T.R., Turley, J.P., Patel, V.L., Zhang, J., 2007. Towards a hybrid method to categorize interruptions and activities in healthcare. Int. J. Med. Inf. 76 (11), 812–820.

Brolin, A., Bäckstrand, G., Thorvald, P., Högberg, D., Case, K., 2012. Kitting as an information source in manual assembly. Adv. Ergon. Manuf. 346–353.

Bäckstrand, G., 2009. Information Flow and Product Quality in Human Based Assembly (Doctoral dissertation). Loughborough University, Loughborough, UK.

Cohen, J., 1994. The earth is round (p<.05). Am. Psychol. 49 (12), 997–1003.

Cooper, A., Reimann, R., Cronin, D., Noessel, C., 2014. About Face: the Essentials of Interaction Design. John Wiley & Sons, Hoboken, NJ.

Coraggio, L., 1990. Deleterious Effects of Intermittent Interruptions on the Task Performance of Knowledge Workers: a Laboratory Investigation (Doctoral dissertation). The University of Arizona. Retrieved from: Interruptions.net/literature/Coraggio-PhD.pdf.

Dix, A., Ramduny, D., Wilkinson, J., 1998. Interaction in the large. Interact. Comput. 11 (1), 9–32.

Donders, A.R.T., van der Heijden, G.J., Stijnen, T., Moons, K.G., 2006. Review: a gentle introduction to imputation of missing values. J. Clin. Epidemiol. 59 (10), 1087–1091.

Drews, F.A., Bakdash, J.Z., 2013. Simulation training in health care. Rev. Hum. Factors Ergon. 8 (1), 191–234.

Floel, A., Poeppel, D., Buffalo, E.A., Braun, A., Wu, C.W.H., Seo, H.J., Stefan, K., Knecht, S., Cohen, L.G., 2004. Prefrontal cortex asymmetry for memory encoding of words and abstract shapes. Cereb. Cortex 14 (4), 404–409.

Freivalds, A., Niebel, B., 2013. Niebel's Methods, Standards, & Work Design. McGraw-Hill Higher Education, New York.

Gillie, T., Broadbent, D., 1989. What makes interruptions disruptive? A study of length, similarity, and complexity. Psychol. Res. 50 (4), 243–250.

Grandhi, S.A., Jones, Q., 2015. Knock, knock! Who's there? Putting the user in control of managing interruptions. Int. J. Hum. Comput. Stud. 79, 35–50.
Groover, M.P., 2010. Fundamentals of Modern Manufacturing: Materials Processes, and Systems. John Wiley & Sons, Hoboken, NJ.
Hawkins, F.H., 1993. Human Factors in Flight. Koninklijke Wöhrmann, Abcoude.
Hermann, M., Pentek, T., Otto, B., 2015. Design Principles for Industrie 4.0 Scenarios: a Literature Review. Working Paper No. 01/2015. Technische Universität Dortmund, Fakultät Maschinenbau, Business Engineering Institute St. Gallen, CH 9008 St. Gallen.
Hollan, J., Hutchins, E., Kirsh, D., 2000. Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans. Comput. Hum. Interact. TOCHI 7 (2), 174–196.
Hutchins, E., 1995. Cognition in the Wild. MIT Press, Cambridge, MA.
Iqbal, S.T., Bailey, B.P., 2008. Effects of intelligent notification management on users and their tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 93–102.
Iqbal, S.T., Horvitz, E., 2007. Disruption and recovery of computing tasks: field study, analysis, and directions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 677–686.
Jackson, S.A., Kleitman, S., Aidman, E., 2014. Low cognitive load and reduced arousal impede practice effects on executive functioning, metacognitive confidence and decision making. PloS One 9 (12), e115689.
Kaptelinin, V., Nardi, B.A., 2006. Acting with Technology: Activity Theory and Interaction Design. MIT Press, Cambridge, MA.
Kirsh, D., 2001. The context of work. Hum. Comput. Interact. 16, 305–322.
Kolbeinsson, A., Thorvald, P., Lindblom, J., 2014. Context aware interruptions: existing research and required research. In: Stanney, Kay, Hale, Kelly S. (Eds.), Advances in Cognitive Engineering and Neuroergonomics, AHFE2014, pp. 260–272.
Kolbeinsson, A., Lindblom, J., 2015. Mind the body: how embodied cognition matters in manufacturing. Procedia Manuf. 3, 5184–5191.
McFarlane, D.C., 1997. Interruption of People in Human-computer Interaction: a General Unifying Definition of Human Interruption and Taxonomy (No. NRL-FR-5510-97-9870). Office of Naval Research, Arlington, VA.
McFarlane, D.C., 1999. Coordinating the interruption of people in human-computer interaction. In. In: Human—Computer Interaction — INTERACT'99, pp. 295–303.
McFarlane, D.C., 2002. Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. Hum. Comput. Interact. 17 (1), 63–139.
McFarlane, D.C., Latorella, K.A., 2002. The scope and importance of human interruption in human-computer interaction design. Hum. Comput. Interact. 17 (1), 1–61.
Mirlacher, T., Palanque, P., Bernhaupt, R., 2012. Engineering animations in user interfaces. In: Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems. ACM, pp. 111–120.
Monk, C.A., Boehm-Davis, D.A., Trafton, J.G., 2002. The attentional costs of interrupting task performance at various stages. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 46(22). SAGE Publications, pp. 1824–1828.
Norman, D.A., 1993. Things that Make Us Smart: Defending Human Attributes in the Age of the Machine. Addison-Wesley, Reading, MA.
Rooksby, J., 2013. Wild in the laboratory: a discussion of plans and situated actions. ACM Trans. Comput. Hum. Interact. TOCHI 20 (3), 19–36.
Rubinstein, J.S., Meyer, D.E., Evans, J.E., 2001. Executive control of cognitive processes in task switching. J. Exp. Psychol. Hum. Percept. Perform. 27 (4), 763.
Rukab, J.A., Johnson-Throop, K.A., Malin, J., Zhang, J., 2004. A framework of interruptions in distributed team environments. Stud. Health Technol. Inf. 107 (2), 1282–1286.
Sanderson, P.M., Grundgeiger, T., 2015. How do interruptions affect clinician performance in healthcare? Negotiating fidelity, control, and potential generalizability in the search for answers. Int. J. Hum. Comput. Stud. 79, 85–96.
Scerbo, M.W., 1998. What's so boring about vigilance? In: Hoffman, R.R., Sherrick, M.F., Warm, J.S. (Eds.), Viewing Psychology as a Whole: the Integrative Science of William N. Dember. American Psychological Association, Washington, DC, pp. 145–166.
Smallman, H.S., St John, M., 2005. Naive realism: misplaced faith in realistic displays. Ergon. Des. Q. Hum. Factors Appl. 13 (3), 6–13.
Speier, C., 1996. The Effects of Task Interruption and Information Presentation on Individual Decision Making (Unpublished doctoral dissertation). Indiana University of Pennsylvania. http://www.interruptions.net/literature/Speier-PhD.pdf.
Speier, C., Vessey, I., Valacich, J.S., 2003. The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. Decis. Sci. 34 (4), 771–797.
Spiekermann, S., Romanow, A., 2008. Attention & Interruption Management for Systems Design—a Research Overview (Disc. paper). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1244622.
Sykes, E.R., 2011. Interruptions in the workplace: a case study to reduce their effects. Int. J. Inf. Manag. 31 (4), 385–394.
Tolmie, P., Crabtree, A., 2008. Deploying research technology in the home. In: Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work. ACM, pp. 639–648.
Tukey, J.W., 1962. The future of data analysis. Ann. Math. Stat. 1–67.
Walter, S.R., Dunsmuir, W.T., Westbrook, J.I., 2015. Studying interruptions and multitasking in situ: the untapped potential of quantitative observational studies. Int. J. Hum. Comput. Stud. 79, 118–125.
Warnock, D., McGee-Lennon, M.R., Brewster, S., 2011. The impact of unwanted multimodal notifications. In: Proceedings of the 13th International Conference on Multimodal Interfaces. ACM, pp. 177–184.
Wickens, C.D., 1992. Engineering Psychology and Human Performance, second ed. Harper Collins, New York.
Womack, J.P., Jones, D.T., 2003. Lean Thinking: Banish Waste and Create Wealth in Your Corporation. Simon and Schuster, London.
Zeigarnik, B., 1927. On finished and unfinished tasks. Psychol. Forsch. 9, 1–85.
Zijlstra, F.R., Roe, R.A., Leonora, A.B., Krediet, I., 1999. Temporal factors in mental work: effects of interrupted activities. J. Occup. Organ. Psychol. 72 (2), 163–185.