Measuring the Effect of Team Diversity and Collective Intelligence in Agile Teams on Software Development Efficiency

Conference Paper · August 2017

CITATIONS
0

READS
305

1 author:

Phil Hennel
University of Cologne
20 PUBLICATIONS **94** CITATIONS

SEE PROFILE

Measuring the Effect of Team Diversity and Collective Intelligence in Agile Teams on Software Development Efficiency

Full Paper

Phil Diegmann
University of Cologne
diegmann@wiso.uni-koeln.de

Abstract

Information system development is largely dependent on social interaction and team work. Team composition, team processes, and behavior among, as well as agile practices used by team members play an important role for the success of information system development projects. Organizational psychology research found team diversity and collective intelligence to be important factors for team performance. In this paper, I evaluate the constructs and items for a model to investigate the effects of team diversity, collective intelligence, interpersonal relationships, and cognitive styles on team efficiency in agile software development. The evaluation is based upon a substantive validity test and a discriminant validity test.

Keywords

Agile Software Development, Diversity, Collective Intelligence, Interpersonal Relationships, Cognitive Styles, Efficiency, Team Composition, IS Development, IS Project Management

Introduction

Agile software development (SD) methods are increasingly popular in the industry (Conboy 2009; Dybå and Dingsøyr 2008; Fitzgerald et al. 2006; Lee and Xia 2010; Williams 2012). With increasing diversity among agile SD (henceforth ASD) teams and an increased need for team management, it has become important to understand the mechanisms of action and the effects present in ASD teams (Lee and Xia 2010; Persson et al. 2012; Sarker et al. 2009).

Team-level research in ASD, on the other hand, is scarce (Lee and Xia 2010), although ASD is mostly conducted in teams and is quintessentially a team effort (Siau et al. 2010). Especially team diversity and collective intelligence (CI) are found to be important predictors of team performance (Phillips et al. 2006; Post 2012; Woolley et al. 2010). Moreover, so far results regarding team-level effects are inconsistent. Some studies suggest that ASD methods work best for highly cohesive (non-diverse) teams (Cao et al. 2009; Fruhling and de Vreede 2006), and cohesiveness could be the main reason for successful SD. Others find that diversity amplifies creativity and communication and therefore contributes to the success of ASD methods (Bear and Woolley 2011; Lee and Xia 2010; Phillips et al. 2006).

Research on teams also has identified a "need to move beyond the simple diversity-affects-performance model in order to think in more complex ways about *how* and *under what conditions* a diversity of expertise in groups might promote or inhibit group effectiveness" (Van Der Vegt and Bunderson 2005, p. 542). Similarly, researchers call for more empirical research on how diversity affects team performance in ASD (Lee and Xia 2010), and on team-level effects in ASD (Conboy 2009; Mangalaraj et al. 2009; McAvoy and Butler 2009; McAvoy et al. 2013).

To shed light on the diversity-performance debate, and to conceptualize this for the domain of ASD, I suggest to build on insights from recent studies on team work to propose a model answering the calls for further conceptualization and investigation. Specifically, ASD teams are involved in very different phases over a project's lifecycle (i.e., planning, implementation, introduction, etc.), of which each demands different specializations and includes different kinds of tasks. Providing deeper insights into benefits and downsides of team-level effects (i.e., diversity and CI) regarding different phases of the ASD lifecycle

would benefit both research and practice. Further, research indicating both increases and decreases in team performance based on team composition (i.e., diversity and CI), giving recommendations on how to compose teams might reduce the number of failed projects and therefore might lead to a decrease in costs. Building upon previous work (Diegmann and Rosenkranz 2016), I evaluate constructs and items on how to measure the impact of CI and team diversity on team efficiency, considering the influences of ASD practices. My evaluation of the substantive validity and discriminant validity is based on a study conducted among researchers and agile software developers. I asked the participants to rate the affinity of each item to the constructs of this model. Based upon the ratings and the feedback given by participants, the measurement items were refined to enable valid testing of this model and to enable future research to utilize agile practices better and, in turn, improve team efficiency.

The remainder of this paper is structured as follows. I give an overview about related work, targeting team effects and SD. Next, I give a brief overview over the underlying model. Finally, I evaluate the proposed measurements and discuss my findings.

Related Work

Information Systems Development and Agile Approaches

ASD methodologies (Cao et al. 2009) emphasize flexibility and autonomy within the team, the overall development process is not planned and scheduled upfront, and progress is made in small iterative phases, and encourage change and constant feedback (Cockburn and Highsmith 2001; Highsmith and Cockburn 2001). Planning thereby becomes a permanent task, and team leadership is established via collaboration and is separated from project lead (Dybå and Dingsøyr 2008). While the team is thus highlighted as the crucial aspect of ASD in practice, extant research has investigated mainly specific and individual or organizational phenomena (Balijepally et al. 2009; Maruping et al. 2009b), and effects regarding whole projects or organizations (Cao et al. 2009; Heeager 2012; Kotlarsky 2007; Mangalaraj et al. 2009). Team-level effects, especially diversity and CI, are covered less so, and existing results are contradictory. Team-research found that cohesive teams are the optimal base for applying agile practices (Cao et al. 2009; Fruhling and de Vreede 2006), while other studies suggest that diversity amplifies creativity and problem solving ability (Bear and Woolley 2011; Lee and Xia 2010; Phillips et al. 2006) and therefore might provide benefits for SD. These inconsistencies are especially important for ASD, as ASD teams rely heavily on efficiency (to respond quickly to requirement changes and being flexible; Conboy 2009) and problem solving ability (to complete complex, non-routine tasks; Lee and Xia 2010).

As SD projects are becoming more distributed and diverse (e.g., Persson et al. 2012; Ramesh et al. 2012; Sarker et al. 2009; Sarker and Sarker 2009), research on ASD is calling for a better understanding of effects of diversity in SD (Lee and Xia 2010). Extant research applied theories of organizational psychology while being focused rather on IT use than on SD (e.g., Gorecki et al. 2008; Nan 2011). While research on teams thus is not completely new to SD research, team composition effects, such as diversity and CI, have not been investigated by SD research yet.

Team Work and Group Performance

Research on team work has focused on outcomes of team performance, before shifting to mediation effects and more general speaking from input-process-output models to cyclic input-mediation-output-input models (for a comprehensive overview, see Ilgen et al. 2005). A notion of teams as complex, context-sensitive, and evolving systems has emerged (Ilgen et al. 2005; Kozlowski and Bell 2003).

In organizational psychology, two concepts have emerged as important predictors of team performance in the last decade: diversity and CI. As regards the first, research over recent decades found contradictions (del Carmen Triana et al. 2014; Joshi and Roh 2009; Milliken and Martins 1996; Phillips et al. 2006; Post 2012; Van Der Vegt and Bunderson 2005). Some studies find a positive relation between diversity and team performance (Bear and Woolley 2011; Phillips et al. 2006; Van Der Vegt and Bunderson 2005), but outlined a dependency on specific contextual circumstances, such as the competitive threat-level (del Carmen Triana et al. 2014), team identification, and climate (Van Der Vegt and Bunderson 2005). Team identification and climate have been found to play an important role in generating positive effects from diversity (Van Der Vegt and Bunderson 2005). Studies which identified a negative effect describe an

overhead of communication and conflict (Ely and Thomas 2001; Leonard et al. 2004; MacMillan et al. 2004).

Scholars differentiate between deep-level (DLD; e.g., education, experiences, values, or believes, also known as job-relevant diversity) and surface-level diversity (SLD; e.g., ethnicity, age, also known as background diversity) (Aggarwal and Woolley 2013; Phillips et al. 2006). These two types act differently: while SLD highlights dissimilarities and encourages sharing of unique information (Phillips et al. 2006), DLD might lead to harmful conflict (Jehn et al. 1999) or facilitate team performance by providing different educational backgrounds and skillsets (Joshi and Roh 2009). Furthermore, diversity can be grouped (Harrison and Klein 2007) into separation (differences in opinions etc.), disparity (differences in socially valued assets), and variety (differences in knowledge etc.).

In regard to CI (Woolley et al. 2010), the average intelligence of team members and the single highest intelligence correlate only weakly with CI, and cohesion, motivation, and satisfaction do not correlate. The two most important factors contributing to CI are social sensitivity and a balanced number of speaking turns per group member (Wooley et al 2010).

While these phenomena have been investigated on their own and mainly in the context of general or occasional teams, SD research has not put these theories together and evaluated these effects in the specific context of ASD teams, although ASD methods rely heavily on team work, composition, communication, and interpersonal relationships (Beck et al. 2001; Lee and Xia 2010; Maruping et al. 2009a; Rosenkranz et al. 2013). Research on team composition and performance has been mostly performed involving students in laboratory conditions, but to my knowledge no studies have empirically investigated effects of team diversity and CI specific to real-world ASD teams.

Research Design

In this paper, I evaluate how to measure the underlying model, which is why I do not explicitly reason for the relationships of the model itself. For a detailed reasoning and theory development, refer to Diegmann and Rosenkranz (2016), which can be summarized as follows. Due to the ever-changing nature of ASD, team efficiency is crucial for project success. As teams become more and more diverse (e.g. due to globally distributed teams), interpersonal relationships play an important role in managing the high potential of conflict and social agile practices (SAPs) help in establishing relationships. Furthermore, CI as a team's general task-solving ability plays an important role in assessing a team's performance (i.e., efficiency). This effect can only take effect if the cognitive styles present match (Diegmann and Rosenkranz 2016). Figure 1 displays the underlying model and Table 1 summarizes the definitions of the constructs used.

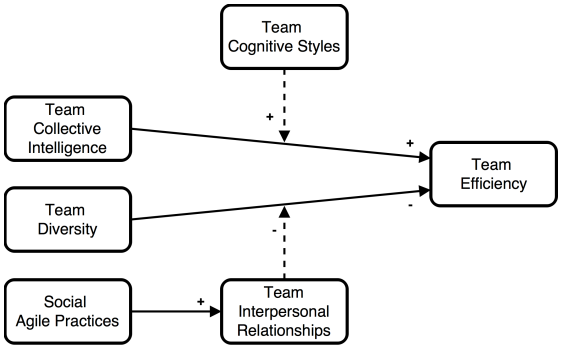

Figure 1. The underlying model according to Diegmann and Rosenkranz (2016)

Construct	Definition
Team Collective Intelligence (CI)	A group's general ability to perform a wide variety of tasks, as a property of the group itself, not of its members. CI is differentiated from the intelligence of individuals on a group level, as average and highest individual intelligence were weaker predictors of group performance and are only moderately correlated with CI. The team-level construct will be composited from individual measurements by following the referent-shift consensus model (Chan 1998).
Team Diversity (DLD)	The differences among team members regarding visible (i.e., surface-level; e.g., race, age) and invisible (i.e., deep-level; e.g., experience, education) characteristics. The team-level construct will

	be composited from individual measurements by following the additive model (Chan 1998).
Team Cognitive Styles (CS)	The way one gathers, processes, and organizes information; typically differentiated into more analytical, process-driven and more connective, holistic thinking. The team-level construct will be composited from individual measurements by following the dispersion model (Chan 1998).
Team Efficiency	A team's ability to complete a project in time and budget, and to respond to requirement changes and incorporate them with minimal time, cost, and resources.
Interpersonal Relationships (IRs)	Relationships and their associated behavior and norms in the workplace; manifestations are business associates, business friends, or personal friends. The team-level construct will be composited from individual measurements by following the direct consensus model (Chan 1998).
Social Agile Practices (SAPs)	ASD practices entailing communication practices or practices aiming at exchanging knowledge and facilitating interpersonal interaction (e.g., daily scrums or pair programming).

Table 1. Definitions of Constructs

To ensure construct validity of the items and to identify ambiguous items, I asked six scholars and five software developers to rate the items affinity to the six constructs. This methodology is similar to the substantive validity and agreement test proposed by Anderson and Gerbing (1991). Based on the rating, I can identify the sorting by the highest rating, while additionally having the possibility to evaluate the discriminant validity by performing a t-test.

I provided a sheet containing definitions for each of the constructs and all measurement items, which are based upon previous literature (see Table 2 and Diegmann and Rosenkranz (2016)). The participants were then asked to read all measurement items for the constructs (see Table 2) and the constructs' definitions (see Table 1) carefully and afterwards rate the affinity of each item with each construct on a scale of 1 ("does not match at all") to 5 ("does match exactly"). Furthermore, I asked the participants to point out all items, which were either worded poorly or ambiguous. Two rounds of evaluations were conducted. During the first round, four researchers, familiar with ASD, were asked to evaluate the items. Based upon their feedback, the items were revised and a second round with two additional researchers and five software developers was conducted. In the next section, I provide details on the evaluation.

Measurement Evaluation & Discussion

Following the reasoning of Anderson and Gerbing (1991), I first calculate the *proportion of substantive agreement*, P_{SA}, which is defined as the extent to which an item reflects its intended construct on a range from -1.00 to 1.00. In addition, I utilize the *substantive validity coefficient*, C_{SV}, as the P_{SA} index does not indicate the extent to which an item might also reflect other, unintended constructs. C_{SV} represents the extent to which respondents assign an item to its posited construct more than to any other construct, ranging from 0.00 to 1.00. For both indices a larger value indicates a greater substantive validity, whereas 0.5 is a recommended threshold for the C_{SV} index (Anderson and Gerbing 1991).

In addition to P_{SA} and C_{SV} I conducted a t-test on the item-affinity-ratings to detect indicators for insufficient discriminant validity. By allowing participants not only to assign an item to a single construct to which it fits best, but also by giving information about possible interference with other constructs by rating higher than 1, I can detect construct-overlapping items. Table 2 displays the results for P_{SA}, and C_{SV} on a per-item level across all rounds and Table 3 displays the results across all rounds on a per-construct level and gives information on the inter-construct discriminant validity results.

Construct References	**Item**	**1st Round**		**2nd Round**	
		P_{SA}	C_{SV}	P_{SA}	C_{SV}
CI Woolley et al. (2010) Engel et al. (2014)	Following, you'll see 36 pairs of eyes. Please indicate, which of the four listed emotions describes the depicted person's emotions best. Please decide as quickly as possible, after reading all options. [a]	.25	.00	.29	-.14
	In the last team meeting, I had less speaking turns than average	.25	.00	.43	.14
	In the last team meeting, someone else had less speaking turns than average	.25	-.50	.29	.00
	In the last team meeting, speaking turns were evenly distributed among	.25	-.25	.57	.29

Construct References	**Item**	**1st Round**		**2nd Round**	
		P_{SA}	C_{SV}	P_{SA}	C_{SV}
	team members				
	In most team meetings, I have less speaking turns than average	.00	-.75	.57	.29
	In most team meetings, someone else has less speaking turns than average	.75	.50	.43	.14
	In most team meetings, speaking turns were evenly distributed among team members	.00	-.75	.57	.43
DLD Phillips et al. (2006) Post (2012)	Please indicate your highest educational level	1	1	.86	.71
	Please indicate your professional experience in the field of agile software development and related areas	.75	.50	.57	.29
	Please indicate all fields in which you gained expertise during education, training or work	1	1	1	1
	Please indicate your level of expertise in the topic of agile software development	.25	-.50	.57	.43
	Did you ever spend a prolonged period abroad or in a different culture?	.75	.50	.86	.71
CS Scott and Bruce (1995)	I double-check my information sources to be sure I have the right facts before making decisions	.75	.50	1	1
	I make decisions in a logical and systematic way	.75	.50	.86	.71
	My decision making requires careful thought	.75	.50	.86	.71
	When making a decision, I consider various options in terms of a specific goal	1	1	.86	.71
	When making decisions, I rely upon my instincts	.75	.50	1	1
	When I make decisions, I tend to rely on my intuition	.75	.50	.86	.71
	I generally make decisions that feel right to me	.75	.50	.86	.71
	When I make a decision, it is more important for me to feel the decision is right than to have a rational reason for it	.75	.50	.86	.71
	When I make a decision, I trust my inner feelings and reactions	.75	.50	1	1
	I often need the assistance of other people when making important decisions	.75	.50	.71	.57
	I rarely make important decisions without consulting other people	.75	.50	.57	.29
	If I have the support of others, it is easier for me to make important decisions	.75	.50	.57	.29
	I use the advice of other people in making my important decisions	.25	.00	.71	.57
	I like to have someone to steer me in the right direction when I am faced with important decisions	.50	.00	.57	.43
	I avoid making important decisions until the pressure is on	.75	.50	1	1
	I postpone decision making whenever possible	.75	.50	.86	.71
	I often procrastinate when it comes to making important decisions	.25	-.50	.71	.57
	I generally make important decisions at the last minute	.75	.50	.86	.71
	I put off making many decisions because thinking about them makes me uneasy	.75	.50	.71	.57
	I generally make snap decisions	.75	.50	1	1
	I often make decisions on the spur of the moment	.75	.50	.71	.57
	I make quick decisions	.50	.00	.86	.71
	I often make impulsive decisions	.25	.00	1	1
	When making decisions, I do what seems natural at the moment	.75	.50	.86	.71

Construct References	**Item**	**1st Round**		**2nd Round**	
		P_{SA}	C_{SV}	P_{SA}	C_{SV}
CS Woolley (2009)	Your team discussed which part of the project the team members will work on next	.00	-.50	.14	-.29
	Your team discussed when each of the subtasks will be completed	.00	-1	.14	-.29
	Your team discussed how the team should divide its time among the various subtasks	.00	-.75	.00	-.71
	Your team discussed how you personally will divide your time among the subtasks	.00	-1	.00	-.71
	Your team discussed what constitutes a "successful performance" on your task	.00	-.75	.29	-.43
	Your team discussed what criteria will be used for evaluating the final product	.00	-.50	.29	-.29
	Your team discussed what information will be helpful to you and your team in working on this project	.00	-.75	.00	-1
	Your team discussed what the final output of your and your team's work will look like	.00	-.75	.00	-.86
Efficiency Lee and Xia (2010) Van Der Vegt and Bunderson (2005)	Meetings always started in time	.75	.50	.71	.57
	Meetings were always finished in time	.75	.50	.86	.71
	Features were always done in time	1	1	1	1
	Bugs were always resolved in time	1	1	1	1
	Overall, during the last sprint, the schedule was not exceeded	.75	.50	.86	.71
	The schedule was always created upfront	1	1	1	1
	The schedule was always known by all team members upfront	1	1	.71	.57
	Progress was tracked throughout the sprint	.75	.50	.86	.71
IRs Madsen and Matook (2010)	Please Indicate the level of communality in your team	1	1	.86	.71
	Please indicate the level of helping behavior in your team	.75	.50	1	1
	Please indicate the type of typical conversations in your team (only essential to only non-essential)	1	1	.57	.29
	Please indicate the amount of self-disclosure in your team	.50	.00	.86	.71
	Please indicate the level of trust in your team	1	1	.86	.71
	Please indicate the level of intimacy in your team	.75	.50	1	1
	Please indicate the level of attention to social needs in your team	1	1	1	1
	Please indicate the level of obligation to return favors in your team	1	1	1	1
	Please indicate the norms present in your team (exchange vs cooperative norms)	.75	.50	.86	.71
SAPs Hummel et al. (2015)	Co-located office space was used	.25	-.50	.57	.29
	Stand-up meetings were held	.75	.50	.71	.57
	An iteration planning meeting was held	1	1	.71	.57
	Pair programming was used	.75	.50	.86	.71
	A sprint retrospective/review was held	.75	.50	.86	.71
	Co-located office space was the main work environment	.75	.50	.57	.29
	Stand-up meetings were held on each work day	1	1	.86	.71
	Iteration planning meetings were held regularly	1	1	1	1
	Pair programming was used rigorously	.75	.50	1	1
	Sprint retrospectives/reviews were held regularly	1	1	.57	.29

Table 2. Measurement Evaluation. Substantive Validity Pre-test per Item
Items with grey background excluded from further use. Cells with orange background mark below-threshold items.
[a] reduced to one item explaining the measurement instead of displaying 36 items consisting of pictures of eyes.

Construct # Items	1st Round				2nd Round			
	P_{SA}	C_{SV}		p-value [b]	P_{SA}	C_{SV}		p-value [b]
CI 6 [b]	.25	-.25	DLD, Efficiency	< .005	.45	.16	DLD, IRs	n.s.
			CS	< .001			CS, Efficiency, SAPs	< .001
			IRs, SAPs	n.s.				
DLD 5	.75	.57	CI	< .05	.80	.65	CI	< .001
			CS	n.s.			CS	< .001
			Efficiency, IRs, SAPs	< .001			Efficiency	< .001
CS [c] 24	.68	.40	CI, DLD, Efficiency, IRs, SAPs	< .001	.83	.71	CI, DLD, Efficiency, IRs, SAPs	< .001
Efficiency 16	.88	.75	CI, DLD, Efficiency, IRs, SAPs	< .001	.88	.79	CI, DLD, Efficiency, IRs, SAPs	< .001
IRs 9	.86	.72	CI, DLD, Efficiency, IRs, SAPs	< .001	.89	.79	CI, DLD, Efficiency, IRs, SAPs	< .001
SAPs 10	.80	.60	CI, DLD, Efficiency, IRs, SAPs	< .001	.77	.61	CI, DLD, Efficiency, IRs, SAPs	< .001

Table 3. Measurement Evaluation. Substantive Validity Pre-test Aggregated per Construct
Cells with orange background mark below-threshold items.
[a] reduced to one item explaining the measurement instead of displaying 36 items consisting of pictures of eyes.
[b] p-value regarding the t-test on significant differences in rating to other constructs, i.e. discriminant validity.
[c] I discarded the scale of Woolley (2009) from further use due to the low rating.

As can be seen in the tables above, not all items and constructs showed sufficient results for P_{SA}, C_{SV}, and the t-test on discriminant validity in the first round. Especially the items for the CS scale by Woolley (2009) did not provide sufficient validity for any item on any measure, which is why these items were dropped from further analysis on a per-construct level. The rather low score for the ASD-experience question in the DLD-section can be explained by the similarity to the SAPs section: some participants interpreted this item as a proxy for general ability to use and implement SAPs rather than a diversity measure among team members. A similar reasoning was given by some participants in regard to the CI-related items: most participants interpreted these items as proxies for IRs or CS. Why single items of the CS- and SAPs-scales had significantly lower scores than the remaining respective items could not be answered clearly by the participants. One suggestion, regarding the single, low-scored SAPs-related item, was that this item was interpreted as a proxy for team efficiency.

On a per-construct-level, CI, DLD, and CS show flaws. While CI showed many low scores on a per-item-level, and therefore clearly shows low per-construct-scores as well, CS showed stronger per-item scores. The non-significant discriminant-validity t-test for the DLD-related items from CS-related items shows, that the low C_{SV} originates in some misunderstandings between these two constructs. All other constructs provide sufficient (i.e., of at least .50) scores for P_{SA} and C_{SV} and highly significant p-values for the t-test on inter-construct discriminant validity.

The results from the second round mostly corroborate the findings of the first round, with some improvements and some deterioration. Most items of the CI-scale and all items of the CS-scale by Woolley (2009) retain insufficient P_{SA} and C_{SV} values. Some items from the CS-scale by Scott and Bruce (1995) show an insufficient C_{SV} value, which, according to the participants, is due to an interpretation of the item as related to very different constructs, varying for each participant. Two items of the DLD scale have received below-threshold ratings for C_{SV}. Both items were related to the diversity in regards to ASD experience – which was interpreted by some participants as a cue to efficiency and SAPs. Similarly, items regarding the usage of co-located office space were interpreted by some participants as a necessity for general tem efficiency and IRPs, which explains the rather low C_{SV} value.

On a per-construct level, all constructs but CI have sufficient P_{SA} and C_{SV} values. Also, aside from CI, all constructs are significantly distinct from each other, providing further support for sufficient discriminant validity. Furthermore, the per-construct ratings improved throughout compared to the first round.

To eradicate the remaining flaws, and to be able to use this measurement to evaluate the underlying model, I propose the following steps. First, an alternative measurement for CI needs to be found, as the current measurement does not provide valid results in this context. Second, following the suggestions of some participants after the second round, the measurement items for IRs should be more detailed and less abstract in wording. Third, and similar to the previous issue, participants suggested to revise the SAPs-related items to clearer differentiate whether or not specific practices were used, and if, how regularly these were used precisely. Having completed these three, yet unresolved, issues, I argue that these measurement items are suitable for validly measuring the underlying model. In my view, the remaining, insufficiently scoring items are quite possibly outliers, resulting from a more complex thought process, involving the provided descriptions of the constructs to the participants. Some participants stated, that after having completed the questionnaire and having had more thoughts about the items and their interrelations, they might have rated some items differently. Partly, participants noted, the descriptions of the constructs were too ambiguous, leaving room for interpretation and therefore overlap with other constructs. This ambiguity was on purpose to not prime the participants by giving away too much information or examples.

Future research could use the measurements outlined in this paper to validate the model proposed by Diegmann and Rosenkranz (2016). In order to do so, future research would have to address some of the limitations of this study and downsides of the measurements, encountered during this evaluation. First and foremost, the CI-measurement does not seem to provide valid results in this context. Future research should therefore develop alternative scales and measurements. Additionally, further refinement of a few of the items presented in Table 2 could also improve the substantive validity coefficient. Future research could provide additional insights with an increased sample size. While this sample consisted of both researchers and practitioners, familiar with the topic of ASD, a sample size of more than eleven might yield slightly different results.

Conclusion

In this paper, I have presented the results of a measurement evaluation for a novel model to investigate the relationships of team diversity, CI, team efficiency and SAPs. I have discussed my findings and provided measurements for the constructs used in this model. This multi-stage pre-test indicates a mixed fit of the items suggested by previous literature. While some items performed well (e.g. IRs), some showed deficits (e.g., CI). I provided reasoning on why that is and how to improve the measurement to enable a valid model testing. Nevertheless, this study has some limitations. The sample partly consisted of researchers, which are familiar with ASD research and therefore might recognize some items and would therefore be biased in their response. Furthermore, this substantive validity is no fail-safe guarantee. The measurement in a large-scale field test could still turn out insufficiently valid or reliable. Next, I will conduct a field study to further validate the items used and to evaluate the underlying model. Upcoming research on this topic will help in utilizing agile practices better and will help to improve team efficiency.

Acknowledgements

This research was supported by a grant from the German Research Foundation (DFG) under record no. RO 3650/8-1.

REFERENCES

Aggarwal, I., and Woolley, A. W. 2013. "Do you see what I see? The effect of members' cognitive styles on team processes and errors in task execution," *Organizational Behavior and Human Decision Processes* (122:1), pp. 92-99.

Anderson, J. C., and Gerbing, D. W. 1991. "Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities," *Journal of Applied Psychology* (76:5), p. 732.

Balijepally, V., Mahapatra, R., Nerur, S., and Price, K. H. 2009. "Are Two Heads Better than One For Software Development? The Productivity Paradox of Pair Programming," *MIS Quarterly* (33:1), pp. 91-119.

Bear, J. B., and Woolley, A. W. 2011. "The role of gender in team collaboration and performance," *Interdisciplinary Science Reviews* (36:2), pp. 146-153.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., and Thomas, D. 2001. "Manifesto for Agile Software Development." *Manifesto for Agile Software Development* Retrieved February 24, 2016, from http://www.agilemanifesto.org/

Cao, L., Mohan, K., Xu, P., and Ramesh, B. 2009. "A framework for adapting agile development methodologies," *European Journal of Information Systems* (18:4), pp. 332-343.

Chan, D. 1998. "Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models," *Journal of applied psychology* (83:2), p. 234.

Cockburn, A., and Highsmith, J. 2001. "Agile Software Development: The People Factor," *IEEE Computer* (34:11), pp. 131-133.

Conboy, K. 2009. "Agility from First Principles: Reconstructing the Concept of Agility in Information Systems Development," *Information Systems Research* (20:3), pp. 329-355.

del Carmen Triana, M., Miller, T. L., and Trzebiatowski, T. M. 2014. "The double-edged nature of board gender diversity: Diversity, firm performance, and the power of women directors as predictors of strategic change," *Organization Science* (25:2), pp. 609-632.

Diegmann, P., and Rosenkranz, C. 2016. "Improving Software Development Efficiency–How Diversity and Collective Intelligence Shape Agile Team Efficiency," in: *International Research Workshop on IT Project Management 2016*. Dublin.

Dybå, T., and Dingsøyr, T. 2008. "Empirical Studies of Agile Software Development: A Systematic Review," *Information and Software Technology* (50:9), pp. 833-859.

Ely, R. J., and Thomas, D. A. 2001. "Cultural Diversity at Work: The Effects of Diversity Perspectives on Work Group Processes and Outcomes," *Administrative Science Quarterly* (46:2), pp. 229-273.

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., and Malone, T. W. 2014. "Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face," *PLoS ONE* (9:12), pp. 1-16.

Fitzgerald, B., Hartnett, G., and Conboy, K. 2006. "Customising agile methods to software practices at Intel Shannon," *European Journal of Information Systems* (15:2).

Fruhling, A., and de Vreede, G.-J. 2006. "Field Experiences with eXtreme Programming: Developing an Emergency Response System," *Journal of Management Information Systems* (22:4), pp. 39-69.

Gorecki, J., Berthon, P., DesAutels, P., Donnellan, B., and Teigland, R. 2008. "The Multinational's Nemesis: The Rise of ICT-Enabled Distributed Collective Intelligence?," in: *ICIS 2008 Proceedings*. pp. 1-6.

Harrison, D. A., and Klein, K. J. 2007. "What's the difference? Diversity constructs as separation, variety, or disparity in organizations," *Academy of management review* (32:4), pp. 1199-1228.

Heeager, L. T. 2012. "Introducing Agile Practices in a Documentation-Driven Software Development Practice: A Case Study," *Journal of Information Technology Case and Application Research* (14:1), pp. 3-24.

Highsmith, J., and Cockburn, A. 2001. "Agile Software Development: The Business of Innovation," *IEEE Computer* (34:9), pp. 120-122.

Hummel, M., Rosenkranz, C., and Holten, R. 2015. "The Role of Social Agile Practices for Direct and Indirect Communication in Information Systems Development Teams," *Communications of the Association for Information Systems* (36, pp. 273-300.

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., and Jundt, D. 2005. "Teams in Organizations: From Input-Process-Output Models to IMOI Models," *Annual Review of Psychology* (56, pp. 517-543.

Jehn, K. A., Northcraft, G. B., and Neale, M. A. 1999. "Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups," *Administrative Science Quarterly* (44:4), pp. 741-763.

Joshi, A., and Roh, H. 2009. "The role of context in work team diversity research: A meta-analytic review," *Academy of Management Journal* (52:3), pp. 599-627.

Kotlarsky, J. 2007. "Re-Engineering at LeCroy Corporation: The Move to Component-Based Systems," *Journal of Information Technology* (22:4), pp. 465-478.

Kozlowski, S. W., and Bell, B. S. 2003. *Work Groups and Teams in Organizations*. London: Wiley.

Lee, G., and Xia, W. 2010. "Toward Agile: An Integrated Analysis of Quantitative and Qualitative Field Data on Software Development Agility," *MIS Quarterly* (34:1), pp. 87-115.

Leonard, J. S., Levine, D. I., and Joshi, A. 2004. "Do birds of a feather shop together? The effects on performance of employees' similarity with one another and with customers," *Journal of Organizational Behavior* (25:6), pp. 731-754.

MacMillan, J., Entin, E. E., and Serfaty, D. 2004. "Communication overhead: The hidden cost of team cognition," in *Team cognition: Understanding the factors that drive process and performance,* E. S. S. M. Fiore (ed.), Washington, DC, US: American Psychological Association, pp. 61-82.

Madsen, S., and Matook, S. 2010. "Conceptualizing Interpersonal Relationships in Agile IS Development," in: *ICIS 2010 Proceedings*.

Mangalaraj, G., Mahapatra, R., and Nerur, S. 2009. "Acceptance of Software Process Innovations - the Case of Extreme Programming," *European Journal of Information Systems* (18:4), pp. 344-354.

Maruping, L. M., Venkatesh, V., and Agarwal, R. 2009a. "A Control Theory Perspective on Agile Methodology Use and Changing User Requirements," *Information Systems Research* (20:3), pp. 377-400.

Maruping, L. M., Zhang, X., and Venkatesh, V. 2009b. "Role of collective ownership and coding standards in coordinating expertise in software project teams," *European Journal of Information Systems* (18:4), pp. 355-371.

McAvoy, J., and Butler, T. 2009. "A Failure to Learn By Software Developers: Inhibiting the Adoption of an Agile Software Development Methodology," *Journal of Information Technology Case and Application Research* (11:1), pp. 23-46.

McAvoy, J., Nagle, T., and Sammon, D. 2013. "Using Mindfulness to Examine ISD Agility," *Information Systems Journal* (23:2), pp. 155-173.

Milliken, F. J., and Martins, L. L. 1996. "Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups," *Academy of Management Review* (21:2), pp. 402-433.

Nan, N. 2011. "Capturing Bottom-Up Information Technology Use Processes: A Complext Adaptive Systems Model," *MIS Quarterly* (35:2), pp. 505-540.

Persson, J. S., Mathiassen, L., and Aaen, I. 2012. "Agile Distributed Software Development: Enacting Control through Media and Context," *Information Systems Journal* (22:6), pp. 411-434.

Phillips, K. W., Northcraft, G. B., and Neale, M. A. 2006. "Surface-level diversity and decision-making in groups: When does deep-level similarity help?," *Group processes & intergroup relations* (9:4), pp. 467-482.

Post, C. 2012. "Deep-level team composition and innovation: The mediating roles of psychological safety and cooperative learning," *Group & Organization Management* (37:5), pp. 555-588.

Ramesh, B., Mohan, K., and Lan, C. 2012. "Ambidexterity in Agile Distributed Development: An Empirical Investigation," *Information Systems Research* (23:2), pp. 323-340.

Rosenkranz, C., Corvera Charaf, M., and Holten, R. 2013. "Language Quality in Requirements Development: Tracing Communication in the Process of Information Systems Development," *Journal of Information Technology* (28:3), pp. 198-223.

Sarker, S., Munson, C. L., Sarker, S., and Chakraborty, S. 2009. "Assessing the Relative Contribution of the Facets of Agility to Distributed Systems Development Success: An Analytic Hierarchy Process Approach," *European Journal of Information Systems* (18:4), pp. 285-299.

Sarker, S., and Sarker, S. 2009. "Exploring Agility in Distributed Information Systems Development Teams: An Interpretive Study in an Offshoring Context," *Information Systems Research* (20:3), pp. 440-462.

Scott, S. G., and Bruce, R. A. 1995. "Decision-Making Style: The Development and Assessment of a New Measure," *Educational and Psychological Measurement* (55:5), pp. 818-831.

Siau, K., Long, Y., and Ling, M. 2010. "Toward a Unified Model of Information Systems Development Success," *Journal of Database Management* (21:1), pp. 80-101.

Van Der Vegt, G. S., and Bunderson, J. S. 2005. "Learning and Performance in Multidisciplinary Teams: The Importance of Collective Team Identification," *Academy of Management Journal* (48:3), pp. 532-547.

Williams, L. 2012. "What Agile Teams Think of Agile Principles," *Communications of the ACM* (55:4), pp. 71-76.

Woolley, A. W. 2009. "Putting first things first: Outcome and process focus in knowledge work teams," *Journal of Organizational Behavior* (30:3), pp. 427-452.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., and Malone, T. W. 2010. "Evidence for a Collective Intelligence Factor in the Performance of Human Groups," *Science* (330:6004), pp. 686-688.