Junior Management Science 1 (2016) 1-300

Junior Management Science

journal homepage: www.jums.academy

Relative Performance Information and Financial Incentives in Multidimensional Task Settings – A Conceptual and Experimental Analysis of Effects on Performance and Attention Towards Task Dimensions

Nadine Eichhorn

Ludwig-Maximilians-Universität München

Abstract

The provision of Relative Performance Information (RPI) is commonly used by firms to increase the performance of their employees. In case employees have to fulfil tasks that involve multiple dimensions, firms have to decide on the dimensionality of RPI and can basically choose between unidimensional and multidimensional Performance Information. I discuss behavioural effects of unidimensional and multidimensional RPI under different compensation schemes and apply a controlled laboratory experiment to empirically test the influence of the dimensionality of RPI on performance and attention towards task dimensions. The study demonstrates that solely the provision of unidimensional RPI improves performance in multidimensional task settings while the effects for multidimensional RPI are small and insignificant. Importantly, both unidimensional and multidimensional RPI imply a negative effect on the performance of participants with a low position in the provided ranking on overall performance. Additionally, the application of performance-based compensation negatively moderates the performance impact of both forms of RPI which though seem to be especially critical for multidimensional RPI. In this regard, multidimensional RPI induces a significant performance decrease compared to unidimensional RPI in the presence of a performance-based contract. The findings provide some indication that a distortion of attention toward the 'quantity' dimension of the applied experimental task may have caused the negative performance effect. Notably, the experimental results do not indicate increased learning effects regarding a task-specific strategy under multidimensional as compared to unidimensional RPI. Overall, no significant differences in attention towards task dimensions between the two forms of RPI can be shown.

Keywords: Relative performance information, Incentive schemes, Multidimensional tasks, Social comparison theory, Competition, Feedback

1. Introduction

Firms frequently provide relative performance information (RPI) to their employees in order to show them how well they are performing in comparison to their peers. The general intuition is that such competitive feedback can increase employee motivation and performance even when it is not linked to monetary rewards.[1] In recent years, RPI has played a fundamental role in firms' attempts to improve the productivity of their workforce through the 'gamification' of tasks.[2] This technique refers to the use of 'game design elements in non-game contexts'.[3] For instance, employees can earn points for the completion of tasks while their respective scores are compared with that of their peers in public performance rankings on virtual leaderboards.[4]

Performance comparisons between employees can be potentially based on several performance metrics since many jobs in today's business world either involve multiple tasks within a single job description or single tasks with multiple dimensions.[5] Exemplary for the latter are customer care jobs in hotel reservation centre where employees may

[1] Cf. Reeves and Read (2013); Silverman (2011); Nordstrom et al. (1991).

[2] Cf. Werbach and Hunter (2012); Reeves and Read (2013); The Economist (2012a).

[3] Deterding et al. (2011), p. 9

[4] Cf. Meder et al. (2013), p. 490.

[5] Cf. Colombino et al. (2012), p. 294; The Economist (2012b); Brown et al. (2014), p. 240.

DOI: http://dx.doi.org/10.5282/jums/v1i1pp100-137

be ranked according to the performance dimensions 'number of bookings handled' as well as 'average value of handled bookings'.[6] As indicated by this case, multidimensional task settings normally confront employees with a trade-off in attention allocation since attention spend on one task dimension cannot simultaneously expended toward another.[7] Clearly, the more time employees in a hotel reservation centre spend on each customer in order to increase the value of the booking, the lower is the number of bookings they can handle in a period of time. This implies that performance rankings on different performance dimensions in multidimensional task settings are interdependent.

Multidimensional tasks, like they are relevant in hotel reservation centre, are present across various industries and functions. In the banking sector, for example, important output dimensions of tellers' performance are 'number of new accounts opened' as well as 'average amount of funds in these accounts'. Also, in the field of software development, online programming contests that are organized by firms like TopCoder Inc. measure the number of problems accurately solved by the developer in a given time period but differentiate point values according to the chosen problem difficulty.[8] An important characteristic of the described multidimensional task settings is that performance in the different task dimensions can be aggregated into one overall performance measure. This performance measure may be for instance the total expected revenue that an employee in a hotel reservation centre generates as well as the total funds in the new accounts opened by a banks' teller or the productivity of a developer in a programming contest. In this respect, various combinations of dimension-specific performance can imply a certain level of overall performance. Therefore, as long as firms are only interested in the maximization of overall performance, they have per se no preference on how employees allocate attention towards the different task dimensions. However, since employees may differ in their abilities to perform various task dimensions, performance can significantly depend on an appropriate attention allocation. Thus, concerning the provision of RPI in multidimensional task setting, potential effects on employees' alignment between attention allocation and individual abilities have to be considered.

With respect to the characterized multidimensional task setting, firms have the opportunity to provide their employees with unidimensional or multidimensional RPI. Whereas unidimensional RPI compares peer performance on the overall performance measure that aggregates the different task dimensions, multidimensional RPI additionally includes comparative information regarding the dimension-specific performance. Based on this differentiation, it arises the question if and how the effects of RPI on performance and attention allocation are determined by its dimensionality. In this regard, prior literature on RPI only offers limited insights into its behavioural effects in multidimensional task environments. Although several experimental studies have generated informative findings on the performance impact of RPI, these mainly relate to task settings with one performance dimension and therefore do not consider a potential influence of RPI on attention allocation.[9]

Recent work of Hannan et al. (2013) indicates that prior findings on the performance effects of RPI cannot be simply generalized to a task environment with multiple performance dimensions. Specifically, the study suggests that employees which have to work on multiple tasks may start shifting their attention to one task when they are provided with task-specific RPI. Since the study is concerned with a setting in which a firm prefers equal effort on the tasks, the results imply that the provision of RPI on multiple tasks can impede overall performance improvements although employees are motivated to exert higher effort. Therewith, the findings emphasize the importance of research on the behavioural effects of RPI in task settings which imply that performance is not only determined by the motivated level of effort but also by the induced attention towards different performance dimensions. However, the work of Hannan et al. (2013) does not apply to the effects of RPI in situations where performance is not a matter of firm-preferred allocations, but is rather determined by the extent to which employees' attention towards task dimensions corresponds to their individual abilities.

Against this background, the goal of this study is to explore how the presence and the dimensionality of RPI affect performance and attention towards task dimensions in multidimensional task settings. In particular, the study aims to compare the effects of unidimensional RPI with that of multidimensional RPI. Following to insights of prior research, the effects are studied under different compensation schemes. Even when financial incentives are not tied to relative performance, they may still have a moderating influence on the behavioural effects of RPI. For instance, Tafkov (2013) indicates that the motivational effects of RPI are greater under an individual performance-based contract than under a flat-wage contract.[10] Moreover, research investigating the design of incentive contracts in multidimensional task environments suggest that financial rewards influence an employee's decision on how to allocate attention between different task dimensions.[11] Therefore, the study compares the effects of unidimensional and multidimensional RPI under flat wage compensation as well as under a performance-based contract which links financial rewards to overall task performance.

In order to address the outlined research interest, chapter 2 firstly analyses behavioural effects of RPI theoretically and presents related empirical findings. Specifically, motivational as well as learning effects of RPI may affect performance and attention allocation and are therefore examined in general

[6]Cf. Berger et al. (2011), p. 5.

[7]Cf. Hannan et al. (2013), p. 554; Holmstrom and Milgrom (1991), p. 25.

[8]Boudreau et al. (2012), p. 4-6; TopCoder (2014).

[9]E.g. Hannan et al. (2008); Azmat and Iriberri (2010); Tran and Zeckhauser (2012); Tafkov (2013).

[10]Cf. Tafkov (2013).

[11]Cf. Holmstrom and Milgrom (1991); Feltham and Xie (1994); Kachelmeier et al. (2008).

and with regard to multidimensional task settings. In addition, the influence of different compensation schemes on the motivational effects of RPI is considered. In the following, chapter 3 derives hypotheses and research questions which are tested within a controlled laboratory experiment. Chapter 4 details the methodical approach by outlining the characteristics of experimental research in general and explaining the applied experimental design as well as the experimental procedures. Thereafter, the results of the conducted laboratory experiment are reported and discussed in chapter 5. Finally, chapter 6 concludes the study and refers to its possible limitations as well as to opportunities for future research.

2. Behavioural effects of Relative Performance Information (RPI)

2.1. Motivational effects of RPI

2.1.1. Motivational effects of RPI in general

From a social-psychological perspective, the provision of RPI relates to the fundamental human social interaction process of social comparison – RPI represents social comparison information which enable a comparison between the self and others.[12] A theory of social comparison processes was first proposed by Festinger (1954) in succession of his theory of informal social communication, which emphasizes the influence of social groups on the formation of opinions. According to social comparison theory, individuals use others not only to fulfil their internal need to evaluate their opinions but also to gain knowledge about their own abilities.[13] In his seminal theory, Festinger (1954) hypothesizes that the internal need for social comparison is motivated by the basic drive to maintain an accurate self-evaluation. Since clear objective standards for such evaluations are not always available, accurate assessments of one's ability rely on the comparison with other persons.[14]

The social comparison processes induced by the provision of RPI are of particular relevance for the current analysis since they are expected to increase an individual's motivation to successfully perform a certain task.[15] Festinger (1954) states that engagement in social comparison evokes a 'drive upward' that implies an 'undirectional push to do better and better'[16] which interrelates with the aim of reducing discrepancies between oneself relative to others. These two forms of pressure manifest in competitive behaviour described as 'action to protect one's superiority'[17] which is associated with higher effort and increased performance. Therefore, related research constitute that social comparison can not only serve self-evaluation but also the motives of self-improvement as well as self-enhancement which reflect the desire to improve oneself and to protect or enhance one's self-esteem.[18]

Against this background and with respect to the large body of research that has evolved on social comparison, a more in-depth analysis is required in order to understand how the provision of RPI can affect motivation through social comparison. First, it is of interest to whom people compare themselves so as to examine requirements and characteristics of social comparison. Originally, the theory of Festinger (1954) poses that social comparison involvement requires that comparison standards are similar regarding the critical dimension. This implies that people do not compare themselves to others that have too divergent abilities.[19] Subsequent research added the factor of the similarity on related attributes such as age or educational background. Related attributes influence the performance on the critical dimension and therefore performance differences could be ascribed to them rather than to the ability on the critical dimension. Overall, the selection of similar standards is required to fulfil the motivational aspect of accurate self-evaluation.[20]

Another more detailed aspect of the question to whom people compare themselves is the direction of comparison which also addresses an individual's level of aspiration in a social comparison process. In this context, the social comparison literature fundamentally distinguishes between upward and downward comparison.[21] Based on his central tenet of a 'drive upward', Festinger (1954) predicts that people compare themselves upwards with others who are of slightly better ability.[22] Wheeler (1966) strengthens the concept of upward comparison by introducing the rank order paradigm, stating that people compare themselves in a performance ranking to those ones who rank most similar yet higher. Concerning the motives of social comparison, upward comparison is directed towards the fulfilment of the need for self-improvement as it 'can provide hope and inspiration'.[23] However, it may also negatively affect self-enhancement and the related motive of self-evaluation maintenance.[24] On the basis of this argument, research by Brickman and Bulman

[12] Cf. Guimond (2006), p. 13; Corcoran et al. (2011), p. 119.
[13] Cf. Festinger (1954); Festinger (1950); Suls and Wheeler (2000).
[14] Cf. Festinger (1954), p. 119.
[15] It should be noted that alternative theoretical approaches exist which may serve to predict a positive effect of RPI on individuals' motivation. Major examples are the following: 1) Based on goal- setting theory (cf. Locke and Latham (1990)), the provision of RPI increases motivation as it induces goal setting (e.g. achieving a high position in a performance ranking. 2) Preference-based theories hypothesize that individuals have competitive preferences and therefore derive utility or disutility from the provision of RPI as it informs them that their performance is superior or inferior to that of their peers (cf. Charness and Rabin (2002); Charness and Grosskopf (2001); Kandel and Lazear (1992)). In this context, also status-seeking preferences are of relevance which imply that status concerns for a higher rank can drive motivation under the provision of RPI (e.g. Congleton (1989); Moldovanu and amd X. Shi (2007)).
[16] Festinger (1954), p. 125.
[17] Festinger (1954), p. 126.
[18] Cf. Wood (1989), p. 232; Brown et al. (2007), pp. 60-61.
[19] Cf. Festinger (1954), p. 120.
[20] Cf. Wheeler (1966); Goethals and Darley (1977); Corcoran et al. (2011).
[21] Cf. Wheeler (1966); Garcia and Tor (2007); Brown et al. (2007).
[22] Cf. Festinger (1954), pp. 126-127; Wheeler (1966), p. 27.
[23] Cf. Suls et al. (2002), p. 161.
[24] Cf. Corcoran et al. (2011), p. 124; Tesser (1988) introduces the self-evaluation maintenance model which hypothesizes that individuals strive for the maintenance or increase of self- evaluation.

(1977) bring up the idea of downward comparison, arguing that comparisons with better performing counterparts are avoided because of the implied threatening to the self-view.[25] Wills (1981) and Taylor et al. (1983) further develop this idea and describe downward comparison as a process which enables people to increase their subjective well-being through comparisons with 'lower status targets'. In contrast to upward comparison, it reflects a process aimed to serve the motives of self-enhancement and self-evaluation maintenance.[26]

Based on the differentiation between upward and downward comparison, the question arises of how the direction of social comparison influences motivation. As past research found out, there is no clear answer to it because the affective outcome of a comparison does not depend on its direction but rather on its salient implication. More precisely, a comparison can imply an assimilative or a contrastive evaluation of the comparison target. Assimilation refers to the belief that one could achieve the same status as the comparison target whereas a contrast in social comparison emphasizes the separation between oneself and the target.[27] Therefore, an assimilation is likely to inspire an individual and to increase motivation in an upward comparison, but can on the other hand cause demotivation in case of a downward comparison. A contrast, however, has the opposite effect in upward comparisons because it may foster negative feelings of the own inferiority and lack of ability. In downward comparison, it engenders positive feelings of relief which can spur the motivation for further performance improvement. Overall, assimilation effects are expected to be predominant in comparisons because large differences between the self and the other typically lead to the cessation of social comparison.[28] Nevertheless, social comparison often entails moderate differences and prior research has identified various factors that influence whether assimilative or contrastive evaluations are implied by a comparison.[29] Commonly, it is stated that assimilation is facilitated by feelings of 'we-ness' through the identification and psychological closeness to the comparison target. Consequently, contrast is promoted by contrary feelings of 'I-ness' which are determined by factors stimulating the salience of a personal self in opposition to a social self.[30]

In his classic paper on the theory of social comparison processes, Festinger (1954) assumes that discrepancies in abilities between oneself and a comparison target result in a 'pressure toward uniformity', which implies assimilation.[31] Building on this assumption, individuals that engage in upward social comparison are expected to behave competitively based on their motivation to improve oneself as well as their desire to maintain a positive self-image and to protect self-esteem.[32] However, research has identified several variables that impact the relationship between upward social comparison and competitive behaviour. Historically, three main variables are emphasized that are closely related to the factors that facilitate assimilation. First, competitive behaviour can be intensified through the commensurability of the comparison which refers to the degree of similarity on related attributes between oneself and the comparison target.[33] Second, the psychological closeness to the comparison counterpart typically fosters competitive behaviour as it confronts an individual with a greater threat to self-evaluation maintenance.[34] Third, the higher the relevance of a performance dimension to an individual's self-definition and therefore to self-evaluation maintenance, the stronger may be the induced competitive behaviour.[35] Whereas all of these variables are highly dependent on the involved individuals, recent literature has also suggested additional situational factors that determine the relationship between social comparison and competitive behaviour. Examples include the number of counterparts in a comparison process but also the proximity of a meaningful standard for comparison, like the number one of a performance ranking or another relevant performance threshold. In this context, competitive behaviour is expected to increase as the number of comparison counterparts decreases and the proximity to a meaningful standard gets higher.[36] Overall, these findings illustrate how the strength of the motivational effects of RPI and the related competitive behaviour are determined by a variety of factors which shape the underlying social comparison.

2.1.2. Motivational effects of RPI in multidimensional task settings

So far, the analysis has applied to the basic social comparison processes and their behavioural consequences accompanying the provision of RPI in general. Building on this foundation, this section examines additional motivational effects of RPI which can arise in multidimensional task settings. As illustrated in the introduction of this study, settings in which employees have to perform tasks which include multiple performance dimensions are ubiquitous in today's business environment. Importantly, they confront employees with a trade-off in attention allocation since attention spend on one task dimension cannot simultaneously expended toward another.[37] In this regard, overall performance can significantly

[25] Cf. Brickman and Bulman (1977).

[26] Cf. Wills (1981), pp. 245-246; Taylor et al. (1983), pp. 20-21; Tesser (1988), p. 5.

[27] Cf. Suls et al. (2002), pp. 161-162; Brown et al. (2007), p. 62.

[28] Cf. Collins (2000), p. 159; Mussweiler (2003), p. 474; Festinger (1954), p. 120.

[29] E.g. Mussweiler (2001); Stapel and Koomen (2005); Collins (2000).

[30] Cf. Stapel and Koomen (2005), p. 1037; Collins (2000), pp. 159-160.; Corcoran et al. (2011), p. 128.

[31] Cf. Festinger (1954), p. 126; Corcoran et al. (2011), p. 121.

[32] Cf. Tesser (1988); Beach and Tesser (1995); Wood (1989); Brown et al. (2007).

[33] Cf. Garcia and Tor (2007), p. 98; cf. Goethals and Darley (1977).

[34] Cf. Tesser (1988), p. 5; Garcia and Tor (2007), p. 98.

[35] Cf. Festinger (1954), p. 130; Tesser (1988), p. 6; Garcia and Tor (2007), p. 98.

[36] Cf. Garcia et al. (2006); Garcia and Tor (2007); Garcia and Tor (2009); Additional situational factors ana- lysed by Poortvliet et al. (2009); Tor et al. (2013).

[37] Cf. Hannan et al. (2013), p. 554; Holmstrom and Milgrom (1991), p. 25.

depend on an appropriate attention allocation based on an individual's dimension-specific abilities. Thus, it is of special interest for this study how motivational effects induced by the use of RPI can influence an employee's decision on how to allocate attention between multiple task dimensions. Particularly, a distinction was made between unidimensional RPI that solely compares overall performance and multidimensional RPI which additionally provides comparative information on dimension-specific performance. Following this differentiation, the analysis first draws on findings on motivational effects based on self-affirmation theory and considers the social-psychological phenomenon of social distinction. It then continues the examination of the motivational impact of RPI in multidimensional task settings by analysing the influence of different competitive attitudes.

Self-affirmation tendencies and the desire for social distinction

Self-affirmation theory is built on the assumption that people are driven by a strong motivation to maintain perceived self-integrity.[38] By definition, self-integrity is 'the sense that, on the whole, one is a good and appropriate person'[39] and its evaluation is shaped by culturally or socially significant standards of integrity. Threats to an individual's perceived self-integrity can arise in different forms, including negative social comparisons which may lower self-evaluation as well as self-esteem and therefore affect self-integrity.[40] Self-affirmation theory focuses on the behavioural responses to such threats, aiming to restore a state of integrity. Overall, people can respond to a threat by the acceptance of the threatening information and corresponding behavioural changes. However, in case the threat relates to an important part of self-integrity, it might lead to defensive psychological adaptions such as the denial or the avoidance of the threat. In this context, self-affirmation theory suggests an alternative to defensive adaptations, namely an indirect psychological adaptation of affirming other resources of self-integrity.[41] This implies that people can compensate a threat to self-integrity in one area by emphasizing successes in other areas.[42]

As already stated, negative social comparison experiences are a potential source for threats toward self-integrity. Yet, based on self-affirmation theory, such a threat may be compensated by a positive social comparison in another domain to the extent that it enables the affirmation of an important self-aspect.[43] This insight has important implications for the motivational effects of RPI in multidimensional task settings. In such settings, multidimensional RPI – as opposed to unidimensional RPI – involves social comparison information which allow people to compare themselves with others on the basis of different dimensions of a task. Thereby, self-threatening social comparison in one dimension can trigger self-affirming tendencies in another dimension.[44] As a consequence, in the presence of multidimensional RPI, people may be motivated to concentrate their attention on a task dimension in which they can proof their competence. This might lead to a distortion of attention in terms of an undesirable shift of attention toward one task dimension and the neglect of other dimensions at the expense of overall performance. In contrast, unidimensional RPI limits the scope of social comparison to the overall performance measure and therefore don't provide the opportunity for self-affirmation in other dimensions.

Another social-psychological phenomenon that appears to be relevant in the analysis of motivational effects of RPI on attention towards task dimensions refers to an individual's desire for social distinction. Such an innate desire to differentiate oneselve from others is argued to be a 'hard-wired trait of human nature'.[45] Past research concerned with the motivational effects of symbolic awards highlights social distinction as a driver in award-seeking behaviour.[46] In the context of the provision of multidimensional RPI which compares people based on different task dimensions, the desire for social distinction can potentially cause motivational effects that are similar to those of a need for self-affirmation. Specifically, it may foster a selective concentration on one or more task dimensions which are seen to be possible sources of distinction.

This reasoning is supported by Snyder and Fromkin (1980) who formulate a theory of uniqueness seeking which is related to the concept of social distinction as it states that individuals strive for a moderate self-distinctiveness.[47] They constitute that similarities with others are perceived as being unpleasant and can thus trigger a motivation to positively reaffirm distinctiveness.[48] However, this assumption contradicts to some extent the prediction of social comparison theory that discrepancies in abilities result in a 'pressure toward uniformity'.[49] Thus, it is important to note that motivational effects based on the drive for social distinction as well as uniqueness seeking are – in contrast to self-affirmation theory – not induced by a negative social comparison experience. Rather, the social comparison facilitated by RPI provides a basis on which people can differentiate themselves from others and it might therefore lead toward a motivation to perform better than the comparison targets. As already described, multidimensional RPI offers various opportunities to satisfy distinction which can result in a strong focus on individual task dimensions and thereby also in a disregard of overall performance. Contrastingly, in the presence of

[38] Cf. Steele (1988); Aronson et al. (1999); Sherman and Cohen (2006).

[39] Sherman and Cohen (2006), p. 185; cf. Steele (1988).

[40] Cf. Tesser et al. (2000) for a detailed analysis of the relationship between the concepts of self-esteem and self-integrity.

[41] Cf. Sherman and Cohen (2006), pp. 186-187; Aronson et al. (1999), pp. 128-129.

[42] Cf. Sherman and Cohen (2006), p. 188.

[43] Cf. Leary and Tangney (2005), p. 287; Sherman and Cohen (2006), p. 190.

[44] Cf. Tesser et al. (2000), p. 1481.

[45] Frey (2007), p. 8.

[46] Cf. Bourdieu (1979); Ginsburgh (2003); Frey (2007).

[47] Cf. Snyder and Fromkin (1980).

[48] Cf. Lynn and Snyder (2002), p. 396; Simsek and Yalicetin (2010), p. 576; Cf. Snyder and Fromkin (1980).

[49] Festinger (1954), p. 126.

unidimensional RPI, people can only positively distinguish themselves based on their overall performance although they act in a multidimensional task setting.

Task-oriented and rivalry-oriented competitiveness

The examination of the motivational impact of RPI in general has already outlined how the presence of RPI enables social comparison which can enforce competitive behaviour described as 'action to protect one's superiority'.[50] On this basis, the consideration of socio-psychological research on competition can provide further insights regarding the motivational influence of RPI on attention towards task dimensions in multidimensional task settings.

Competition, defined in situational terms by Deutsch (1949), describes a setting in which the goal attainment of separate participants is negatively correlated, implying that a participant can only achieve his goal if the related others fail in achieving their goals.[51] With regard to this definition, RPI induces competition since it facilitates social comparison which motivates individuals to achieve a superior position in a performance ranking as compared to their counterparts. Based on the early research on competition[52], its psychological consequences and behavioural effects have been extensively studied. An important distinction underlying many of these analyses applies to the different forms of competiveness that it evokes.[53] According to Griffin-Pierson (1990), competitiveness can be defined as an achievement motive or component of achievement motivation.[54] So far, this study has not yet considered different types of competitiveness underlying the competitive behaviour which results from the provision of RPI. However, they might be a determinant of attention towards task dimensions as part of an individual's competitive behaviour. Therefore, the two main types of competitiveness which can arise from competition are discussed briefly.

The first type of competitiveness is characterized by its focus on a performance goal and the related desire to successfully perform a task and 'to be the best one can be'.[55] Various authors have described this competitiveness perspective but have labelled it differently, mainly as 'goal competitiveness'[56], 'personal development competitiveness'[57] or as 'task-orientated competitiveness'.[58] Fundamentally, this type of competitiveness refers to a traditional understanding of competition causing 'behaviour oriented toward a goal in which the other competitors for the goal are secondary'.[59] This implies that the social comparison processes induced by RPI motivate an individual primarily to increase task-related effort as well as effort directed toward learning in order to improve itself.[60] In the following, the study refers to this achievement motive as task-oriented competitiveness.

On the contrary, studies on competition also emphasize a second type of competitiveness that has its focus not on a goal but on others and is thereby called 'interpersonal competitiveness'[61], 'hypercompetitiveness'[62] or 'other-referenced competitiveness'[63]. Inherent to this type of competitiveness is an individual's desire to win and to avoid losing, accompanied with the enjoyment of interpersonal competition.[64] Thus, competition that involves this achievement motive is also framed as rivalry which is partly treated as a separate concept in past literature.[65] From the social comparison perspective, the presence of RPI can thus emphasize a motivation that is mainly about outperforming competitors as well as achieving a higher rank and not in principle on mastering a task as good as possible.[66] In this respect, an individual's needs for self-enhancement and for maintaining a positive self-image dominate the motive of self-improvement. The study labels this competiveness perspective as rivalry-oriented competitiveness.

The characterization of the two forms of competitiveness suggest that they might cause different behavioural responses to RPI in situations that offer people the opportunity to increase their position in a ranking by other means than improving overall performance. Exemplary, prior research argues that rivalry-oriented competitiveness can promote dysfunctional effects of competition like reduced reporting honesty as well as sabotage of comparison counterparts.[67] In multidimensional task settings, it might result in unintended behaviour with regard to people's attention towards the different task dimensions. In particular, multidimensional RPI enforces competition on several task dimensions which provides an individual with the chance to outrank their competitors in at least one dimension without necessarily achieving a better performance on the multidimensional task overall. Therefore, to the extent that it evokes rivalry-oriented instead of task-oriented competitiveness, multidimensional RPI might drive the motivation to focus on winning one of the competitive rankings disregarding the consequences for performance in other dimensions. This might imply a distortion of attention which can be in principle compared to the described behavioural consequences of self-affirmation and social distinction. Comparably, there is no distortion effect expected in case of the provision of unidimensional RPI because it indicates that ranking improvements can be only achieved through enhancements related

[50] Festinger (1954), p. 126.
[51] Cf. Deutsch (1949).
[52] E.g. Triplett (1898); Deutsch (1949); Johnson and Johnson (1989); May and Doob (1937).
[53] Cf. Tassi and Schneider (1997), p. 1560; Griffin-Pierson (1990), p. 108; Ryckman et al. (1996), p. 375.
[54] Cf. Griffin-Pierson (1990), p. 108.
[55] Griffin-Pierson (1990), p. 108.
[56] Cf. Griffin-Pierson (1990); Morey and Gerber (1995).
[57] Cf. Ryckman et al. (1996).
[58] Cf. Tassi and Schneider (1997).
[59] Mead (1937), p. 8.
[60] Cf. Tassi and Schneider (1997), p. 1559; Ryckman et al. (1996), p. 375.
[61] Cf. Griffin-Pierson (1990); Morey and Gerber (1995).
[62] Cf. Ryckman et al. (1990); Ryckman et al. (1996).
[63] Cf. Tassi and Schneider (1997).
[64] Cf. Ryckman et al. (1990), p. 630; Griffin-Pierson (1990), pp. 108-109.
[65] Cf. Mead (1937), p. 8; Kilduff et al. (2010), pp. 945-946; Tassi and Schneider (1997), p. 1558.
[66] Cf. Hibbard and Buhrmester (2010), p. 413; Kilduff et al. (2010), p. 961.
[67] Cf. Brown et al. (2014); Schreck (2014); Charness et al. (2013).

toward the aggregated measure of task performance.

2.2. Learning effects of RPI

2.2.1. Learning effects of RPI in general

The preceding analysis on the motivational effects of RPI is built on the finding that RPI represents an important type of information which facilitate the self-evaluation of abilities in the absence of an objective standard. Therefore, the provision of RPI serves an informational function which not only has implications for an individual's motivation but also affects its learning on how to improve performance.

According to the feedback intervention theory proposed by Kluger and Denisi (1996)[68], RPI is expected to trigger learning processes as it informs an individual about a discrepancy between its own performance and a performance standard.[69] Particularly, such a feedback-standard comparison may stimulate the search for a task-specific plan or the development of a new task-specific strategy in case an increase in effort as a 'universal strategy' proves not to be successful.[70] The learning process involves the generation and re-evaluation of working hypotheses about appropriate strategies with which one's performance can be improved. These hypotheses, in turn, also function as standards in the evaluation of the success of changed behaviour.[71] In this respect, past research on social comparison emphasizes the role of upward comparison in providing information on how to advance and make progress regarding a particular task.[72] Notably, upward comparative information can tell people 'what they would have to achieve to move up in the ability distribution'[73] and the superior counterparts might serve as role models in the learning process.[74]

However, Kluger and Denisi (1996) also point out that the effectiveness of a feedback intervention is determined by the extent to which it directs information to enhanced self-efficacy.[75] Self-efficacy can be defined as an individual's belief in its own ability to influence specific situations that affect its live.[76] The various psychological consequences of self-efficacy beliefs include effects on attentional and thinking processes. Bandura and Wood (1989) state that strong self-efficacy directs people's attention toward problem solving and the development of task strategies. On the other hand, self-doubts of efficacy may strengthen concerns on personal deficiencies as well as the concentration on possible adverse consequences of failures.[77] Consequently, feedback can not only affect an individual's learning through the information it provides about how to develop further, but also through its effect on the cognitive processes related toward the effective use of this information.

The relationship between the presence of RPI and an individual's self-efficacy beliefs depend on whether the involved social comparison implies assimilative or contrastive evaluations.[78] As described earlier, assimilation refers to an identification with the comparison target. Therefore it is suggested to raise self-efficacy in upward comparisons as it demonstrates the attainability of certain goals. On the contrary, a comparison with other from a contrastive perspective rather leads to frustration and resentment associated with lower self-efficacy.[79] Following this logic, assimilation and contrast oppositely affect perceived self-efficacy in downward comparisons. Thereby, it can be concluded that both upward assimilation as well as downward contrast in social comparison processes might increase the learning from RPI because of their positive effect on self-efficacy beliefs.

2.2.2. Learning effects of RPI in multidimensional task settings

In multidimensional task settings, learning can be related toward the improvement of dimension-specific performance as well as the refinement of the allocation of attention towards the different task dimensions. Focussing on the latter, the following analysis aims to answer the question to what extent RPI can help individuals to find an appropriate strategy on how to divide attention based on their dimension-specific abilities. For this purpose, the effects are differentiated between unidimensional and multidimensional RPI consistent with the analysis on the motivational effects of RPI.

The previous section has outlined that RPI can facilitate learning as it allows a feedback-standard comparison which is necessary for the generation and re-evaluation of working hypotheses regarding performance enhancing task-specific strategies.[80] Applied to a multidimensional task setting, the development of a strategy on the allocation of attention requires the formation of hypotheses regarding its relationship with dimension-specific performance and overall performance. In the presence of unidimensional RPI, an individual can only base these hypotheses on comparative information regarding the overall performance measure. However, this implies that an individual cannot disentangle the drivers of the performance of superior and inferior others. Especially, it does not have the opportunity to compare its strategy along with its dimension-specific abilities to that of its competitors. Therefore, an individual might generate a multitude of hypotheses since unidimensional RPI does not provide enough

[68] The feedback intervention theory by Kluger and Denisi (1996) integrates several theories and re- search paradigms on the concept of feedback. Feedback interventions are defined as 'actions taken by (an) external agent (s) to provide information regarding some aspect(s) of one's task performance' (Kluger and Denisi (1996), p. 255).

[69] Cf. Kluger and Denisi (1996), pp. 259-260.

[70] Cf. Kluger and Denisi (1996), p. 263; cf. Wood and Locke (1990).

[71] Cf. Kluger and Denisi (1996), p. 263.

[72] Cf. Corcoran et al. (2011), p. 124; Collins (1996), p. 52

[73] Collins (1996), p. 52.

[74] Cf. Lockwood et al. (2002), p. 854; Corcoran et al. (2011), p. 124; Suls et al. (2002), p. 161.

[75] Cf. Kluger and Denisi (1996), p. 276.

[76] Cf. Bandura (1977b) has established the self-efficacy theory in succession of the social learning theory (Bandura (1977a)).

[77] Cf. Bandura/Wood (1989), p. 806.

[78] Cf. Carmona et al. (2008); Suls et al. (2002).

[79] Cf. Bandura and Jourden (1991); Lockwood et al. (2002); Carmona et al. (2008).

[80] Cf. Kluger and Denisi (1996), p. 263.

information so that erroneous hypotheses can be rejected.[81] By contrast, multidimensional RPI informs an individual on how its dimension-specific and its overall performance compare with that of its counterparts. This provides the basis for the evaluation of hypotheses on the relation between dimension-specific performance, allocation of attention and overall performance. From that, inferences can be drawn with regard to the relative success of certain strategies. Consequently, multidimensional RPI appear to be more helpful than unidimensional RPI in supporting an individual's learning on how to allocate attention between task dimensions in order to improve performance.

Another concern of the preceding analysis on the learning effects of RPI in general has been the influence of RPI on self-efficacy. In general, RPI that implies upward assimilation or downward contrast can increase self-efficacy beliefs.[82] However, it arises the question if unidimensional and multidimensional RPI might differently affect self-efficacy and therewith also the effective use of the provided information for the development of task-specific strategies. Against this background, two determinants of self-efficacy can be identified that may vary depending on the informational value of RPI. Firstly, as already indicated before, self-efficacy can be enhanced by vicarious experience to the degree that it shows an individual that others can overcome difficulties in performing a task.[83] Vicarious experience rely on social comparison and has a higher impact on efficacy as it provides an individual with information on the performance strategies of superior ones.[84] In this regard, multidimensional RPI potentially can raise self-efficacy in a greater manner than unidimensional RPI. The second relevant determinant of self-efficacy is task complexity which is suggested to lower efficacy beliefs. An important component of complexity is an individual's uncertainty on its abilities for performance in situations in which the precise skills required are unclear.[85] Here, multidimensional RPI is expected to contribute to the reduction of complexity by providing social comparison information that links dimension-specific and overall performance and can therefore help to lower uncertainty on skill requirements. Overall, it appears that multidimensional RPI is more beneficial for learning processes on task strategies than unidimensional RPI not only because of the information it provides about how to allocate attention appropriately, but also through its predicted effects on self-efficacy.

2.3. Influence of financial incentives on motivational effects of RPI

2.3.1. Influence of financial incentives in general

From a standard economic perspective, the provision of RPI should not affect an individual's motivation and performance as long as compensation is not based on relative performance.[86] Therefore, the study has relied on social-psychological theories in analysing the motivational effects of RPI. This section now serves to integrate findings of economic theories in order to examine the influence of financial incentives on the predicted effects of RPI. Firstly, insights from agency theory are provided so as to analyse how financial rewards influence motivation. Afterwards, it is specifically considered how individual performance-based compensation might affect the motivational impact of RPI.

To begin with, conventional economic theories predict that a performance-based compensation contract compared to a flat wage contract motivates individuals to exert higher effort and to achieve better performance.[87] Agency theory, in particular, derives this prediction from the fundamental assumptions that individuals are fully rational and solely motivated by self-interest. On this basis, individuals are presumed to strive for a maximization of their expected utility which is described by a utility function that represents preferences for increases in wealth and leisure. Since performing a certain task involves a reduction in leisure, individuals are expected to not exert effort on that task as long as it does not contribute to their economic well-being.[88] Typically, a flat wage contract does not satisfy this condition because it implies that compensation is not linked to an individual's performance. In this context, the principal-agent branch of the agency theory is concerned with the optimal design of compensation contracts in case an individual (principal) engages another individual (agent) to perform a task on its behalf while it cannot fully observe the exerted effort. Notably, as principal-agent models commonly assume that the agents are strictly risk-averse, compensation contracts have to appropriately balance the need for motivating effort against the need for risk-sharing.[89]

Principal-agent theory refers to RPI in the context of Relative Performance Evaluation (RPE) which implies that an individual's compensation is tied to its performance relative to that of its peers. Holmstrom (1982) argues that RPE induces higher effort from risk-averse agents than individual performance-based compensation in case the agents face some common uncertainty. Since RPE allows to filter out common uncertainty from an agent's performance, it reduces an agent's exposure to 'systematic risk' without lowering its incentives.[90] Originating from this idea, past research discusses tournament compensation schemes as a specific type

[81]Cf. Kluger and Denisi (1996), pp. 263-264; Hattie and Timperley (2007), p. 93; Castellan and Swaine (1977).

[82]Cf. Bandura and Jourden (1991); Lockwood et al. (2002); Carmona et al. (2008).

[83]Cf. Bandura (1977b), p. 197.

[84]Cf. Bandura (1977b), p. 197; Gist and Mitchell (1992), p. 194.

[85]Cf. Gist and Mitchell (1992), p. 194; Cervone et al. (1986), p. 492.

[86]Cf. Holmstrom (1982); Kandel and Lazear (1992).

[87]Cf. Bonner and Sprinkle (2002) for an overview of relevant theories and of research findings on variables that moderate the relationship between financial incentives and performance (e.g. person variables, task variables, environmental variables etc.). It should be noted that conflicting social-psychological theories exist which predict negative effects of performance-based compensation, e.g. Deci and Ryan (1985) argue in their self-determination theory that financial incentives may crowd out intrinsic motivation and therefore decrease effort and performance.

[88]Cf. Baiman (1982); Baiman (1990); Eisenhardt (1989).

[89]Cf. Baiman (1990), pp. 342-343.

[90]Cf. Holmstrom (1982), p. 334-337; Antle and Smith (1986), p. 2.

of RPE. Comparably to RPI that is provided in the form of performance rankings, tournament schemes link compensation to an individual's ordinal rank in a tournament.[91]

As already mentioned, based on economic theories, RPI itself should have no effect when compensation is linked to individual performance instead of relative performance.[92] Nevertheless, considering both economic and social-psychological theories, prior research suggests that the outlined positive incentive effect of individual performance-based compensation and the predicted motivational impact of RPI have an accumulative influence on an individual's performance.[93] However, it appears to be unclear if the motivational effects of financial incentives and RPI are only additive or if there is an interaction mechanism between them. Bonner and Sprinkle (2002) conclude from a review of past literature that incentive and feedback effects in general do not interact but are typically 'independent and additive so that there is no simple two-way interaction'[94]. Yet, in a recent study, Tafkov (2013) aims to develop theory that specifically explains how individual performance-based compensation moderates the effect of RPI on performance.[95] With regard to social comparison theory, Tafkov (2013) hypothesizes that RPI is more useful for social comparison purposes under individual performance-based compensation compared to a flat wage contract. The underlying reasoning is that performance-based compensation motivates higher effort levels which results in a clearer relationship between task performance and ability. Therefore, individuals are more likely to attribute performance differences to differences in abilities which is expected to result in greater social comparison involvement and in turn to increase the motivational effect of RPI.[96] Although the study finds empirical support for this prediction, additional theoretical analyses seem to be required in order to substantiate the socio-psychological underpinnings of such an interaction mechanism. Hence, overall, there remains uncertainty whether the motivational effects of financial incentives and RPI are additive or they reinforce each other through an interaction effect.

2.3.2. Influence of financial incentives in multidimensional task settings

The analysis of the motivational impact of RPI in multidimensional task environments has shown that RPI cannot only affect effort but also peoples' attention towards task dimensions. Comparably, economic theory considers the complexities inherent in providing appropriate financial incentives in multitask and multidimensional task environments. Past literature commonly labels attention towards different tasks or task dimensions as effort allocation. In particular, agency-theoretic models are concerned with effort distortion effects which arise when compensation contracts are incomplete because multiple tasks of a job, or multiple dimensions of a task, vary in their measurability.[97] In this context, research also investigates how social pressure or social preferences like fairness influence effort distortion effects in the presence of incomplete compensation contracts.[98] However, since this study focus on settings in which all task dimensions are measurable, this stream of literature is not examined in further detail. Rather, it is of relevance how financial rewards based on an individual's overall performance influence the expected motivational effects of RPI in multidimensional task settings.

The previous section has shown that financial incentives and RPI both increase an individual's motivation to successfully perform a task which might result in an additive performance effect or even a positive interaction. In a multidimensional task setting, yet, the induced motivations might conflict with each other to the extent that they are directed towards different task dimensions. Particularly, financial incentives based on overall performance are expected to motivate an appropriate allocation of attention between different task dimensions.[99] That is, individuals should be motivated to allocate attention according to their hypothesis of how to increase overall performance based on their dimension-specific abilities. This complies with the motivational effects of unidimensional RPI but can conflict with a motivation to focus entirely on one task dimension which is potentially caused by multidimensional RPI. The latter refers to a potential distortion of attention toward a dimension in which an individual can prove its competence induced by self-affirmation tendencies, the desire for social distinction or a rivalry-oriented competitive attitude. In this context, individuals potentially face a trade-off in performance goals when provided with multidimensional RPI while compensation is tied to overall performance. It is important to note that individual performance-based compensation itself is not expected to affect self-affirmation, social distinction or rivalry as long as the monetary rewards received by competitors are not or at least not precisely known.[100] Hence, an individual's behavioural response toward such a trade-off is likely to be determined by the strength of the countervailing effects.

The motivational effects of incentive compensation are suggested to be intensified by the magnitude of a certain payoff as well as the probability of achieving it.[101] The strength of the mentioned socio-psychological processes, however, is fundamentally based on the underlying social comparison. As noted earlier, the commensurability of a comparison target, the psychological closeness toward it and the relevance

[91] Cf. Lazear and Rosen (1981); Nalebuff and Stiglitz (1983).
[92] Cf. Holmstrom (1982); Kandel and Lazear (1992).
[93] Cf. Bonner and Sprinkle (2002); Kluger and Denisi (1996); Chung and Vickery (1976).
[94] Bonner and Sprinkle (2002), p. 329.
[95] Cf. Tafkov (2013).
[96] Cf. Tafkov (2013), pp. 330-331; cf. Festinger (1954).
[97] E.g. Holmstrom and Milgrom (1991); Datar et al. (2001); Feltham and Xie (1994); Baker (2002).
[98] For example Fehr and Schmidt (2004); Brüggen and Moers (2007); Schmidt (2011).
[99] Cf. Datar et al. (2001); Baiman (1990); Prendergast (1999).
[100] Cf. Frey (2007), p. 8.; cf. Sherman and Cohen (2006); Kilduff et al. (2010).
[101] Cf. Bonner and Sprinkle (2002); Baiman (1982); Baiman (1990).

of a performance dimension reinforce the psychological effects of social comparison.[102] It is worth mentioning that rivalry-oriented competitiveness, which may be triggered by these factors, can motivate an individual to apply a task strategy that improve its position in a ranking at the cost of its economic wealth.[103] Past research associates such economically irrational behaviour with a psychological state of 'competitive arousal' which results from rivalry.[104] It implies that an individual's desire to outperform its competitors in a specific dimension of a multidimensional task can dominate its motivation to increase its wealth through a concentration on overall performance. Yet, in sum, no general conclusion can be drawn with regard to the behavioural outcomes resulting from the potentially conflicting motivations induced by financial incentives and RPI.

2.4. Empirical findings of related studies

Various experimental studies have generated empirical findings on the performance impact of RPI resulting from its behavioural effects. Based on the preceding theoretical analysis, this section reviews existing evidence on the effect of RPI on performance in general and presents empirical insights on its interaction with financial incentives. Afterwards, it outlines the findings of the few studies concerned with the effects of RPI in multitask environments since evidence from multidimensional task settings do not exist yet.

At first, studies are considered which explore performance effects of RPI based on field data. Here, recent work of Blanes i Vidal and Nossol (2011) is of relevance as it applies a data set from a firm-level quasi experiment in order to study the effect of feedback on relative performance under a piece rate pay scheme.[105] In particular, the researchers review personnel records of a German wholesale and retail organization which started to privately communicate to the workers of its main warehouse how they rank in the productivity/wage per hour distribution. The findings suggest a long-lasting 6.8 percent increase in workers' productivity following the start of the provision of the ranking information. Additionally, Blanes i Vidal and Nossol (2011) emphasize that this effect can be clearly attributed to the motivational effects of social comparison since the institutional settings implies that the ranking did not have any compensation or career consequences. In another context, research by Azmat and Iriberri (2010) also proves positive performance effects of RPI even though individuals are rewarded according to their absolute performance.[106] The study exploits data from a natural experiment that took place in a Spanish high school in which students received for only one year private information on their performance relative to the average performance in their class. The analysis shows a 5.0 percent increase in the student's grades in reaction to the provision of this information. Notably, the positive effect was significant for the whole distribution and disappeared once the information was removed.

Additional insights on the effects of RPI, however, can be drawn from the results of a field experiment analysed by Tran and Zeckhauser (2012).[107] The researchers also investigate the impact of RPI on academic performance but distinguish the effects between private and public disclosure of ranking information. Specifically, in the course of the field experiment conducted at a University in Vietnam, students enrolled in a one-semester English course were divided into three groups and either received private, public or no information on how their performance ranked relative to their peers after biweekly progress tests. Although the performance rankings brought no direct tangible benefits, the analysis reveals that the provision of RPI resulted in a higher performance of the treatment groups compared to the control group in a final test at the end of the course. Interestingly, the findings imply that the performance increases compared to the control group were around 45 percent higher for those students who received their rank publicly instead of privately.

Further evidence on the performance impact of various types of RPI are provided by laboratory experiments which also offer insights on the moderating effects of different compensation schemes. To begin with, Tafkov (2013) proves in a laboratory environment that both private and public RPI positively influence performance but that the effects are magnified when RPI is public.[108] Those findings are based on a real-effort experiment in which participants had to solve multiple-choice multiplication problems for nine independent rounds without any outside aid. At the end of rounds three, six, and nine, participants in the RPI-condition were informed about the rank of their performance within a group of five participants. In this setting, Tafkov (2013) not only manipulate the type of RPI between subjects but also test the effects of RPI under a flat-wage as well as an individual-performance-based contract. As already noted in section 2.3.1, the results support the hypothesis that the positive performance effects of RPI are stronger under performance-based compensation as it contributes to a clearer relationship between task performance and ability.

In contrast to the insights of Tafkov (2013), an experiment conducted by Murthy (2011) does not reveal significant interaction effects between RPI and different compensation schemes.[109] However, the study compares the effects of RPI under a piece-rate scheme and under two target-based schemes but does not consider a flat wage contract. The target-based schemes consisted of a budget-based scheme and quota-based scheme which both represent performance-contingent compensation contracts. Similar to Tafkov (2013), participants of the experiment of Murthy (2011) had to solve a real-effort task which, yet, involved three periods of decoding numbers into letters. After every

[102] Cf. Tesser (1988); Garcia et al. (2006); Goethals and Darley (1977).
[103] Cf. Malhotra (2010); Ku et al. (2005); Garcia et al. (2006).
[104] Cf. Malhotra (2010), p.140; Ku et al. (2005), pp. 92-93.
[105] Cf. Blanes i Vidal and Nossol (2011).
[106] Cf. Azmat and Iriberri (2010).
[107] Cf. Tran and Zeckhauser (2012).
[108] Cf. Tafkov (2013).
[109] Cf. Murthy (2011).

round, they were publicly informed about their position in a ranking of all 20 participants of the session. Overall, the results reveal a significant performance effect of RPI which is remarkably highest for participants in the bottom third of the performance distribution. Although no interaction effects are found between the provision of RPI and the different compensation schemes, the application of a quota-based contract led to significant performance enhancements which indicate an additive effect with RPI.

The review of empirical findings regarding the interaction between RPI and different compensation contracts can be expanded by considering two experimental studies which compare the performance impact of RPI under an individual incentive scheme and under a tournament scheme. First, a study of Hannan et al. (2008) examines the interaction effects between RPI and the two compensation contracts while varying the precision of RPI.[110] In particular, participants in the RPI-condition of the experiment were either provided with 'coarse' or with 'fine' feedback referring to whether RPI showed the half or the percentile of a performance distribution in which someone ranked relative to others. The task that was applied in the experiment required participants to act as managers and to make output decisions for a single product over 12 trials. The amount of profit points which they earned through their decisions were influenced by a randomly determined state of nature which induced a common error in the setting. Participants had to develop a strategy in order to effectively perform the task and they could evaluate its relative success through the RPI which was privately provided after every fourth trial. Overall, the experimental results show a disordinal interaction between RPI and the compensation scheme since RPI improved mean performance under an individual scheme regardless of its precision but negatively affected it under a tournament scheme if it was precise. The analysis indicates that this mean performance deterioration was not caused by reduced effort but rather by the adoption of ineffective strategies of those who performed below the 80th percentile in the first half of the tournament.

In addition to the insights of Hannan et al. (2008), a second study published by Eriksson et al. (2009) renders evidence on how the frequency of RPI impacts performance under the different incentive schemes.[111] In the experiment, participants were matched in pairs and had to add sets of two-digit numbers individually while they were either continuously informed about the score of their competitor or received that information halfway during the working period. Surprisingly, the results here do not comply with previous findings since they suggest that RPI does not improve performance, regardless of its frequency and the pay scheme. Instead, it is shown that the continuous provision of RPI causes subjects in both pay schemes to make more mistakes compared to when they receive no RPI. Notably, the study does not find an interaction effect between RPI and the compensation scheme. However, supplementary experimental sessions reveal that the lack of positive effects of RPI in the individual pay scheme may be due to subjects already exerting maximum effort in the absence of RPI which is why there is little room for improvements.

The presented empirical studies generate informative insights on the effects of RPI and mainly demonstrate that the provision of RPI leads to significant improvements in performance. This gives support to the expected positive influence of RPI on an individual's motivation and learning. However, the findings solely relate to unidimensional task settings with one relevant performance dimension. A first attempt to examine the generalizability of the positive performance impact of RPI to settings with multiple performance dimensions is made by Hannan et al. (2013).[112]

In particular, this study aims to capture an organizational setting in which employees have to fulfil multiple tasks with diminishing marginal returns whereby the firm has specific preferences over the allocation of effort across those tasks. Therefore, participants in the experiment had to work on both a math task and a verbal task while their compensation was not based on performance but on an equal effort allocation. RPI were provided after each of four rounds either in public or in private. Importantly, it informed participants on how they ranked in each of the two tasks among a group of five contestants. With regard to the preceding theoretical analyses, this form of RPI might induce a distortion effect in a multitask environment based on an individual's tendency to seek self-affirmation, its desire for social distinction or a rivalry-oriented competitive attitude. Accordingly, the study shows that the presence of RPI increased motivation but also induced an effort distortion effect as participants concentrated on one task while reducing their performance on the other task. Both behavioural effects were magnified under public RPI. In this context, however, the induced effort distortion was only detrimental to performance when RPI were public but did not negatively affect it under private RPI.

In a recent working paper of Hannan et al. (2014)[113], the experimental instrument used in Hannan et al. (2013) is adapted in order to examine how the effort distortion effects of public RPI depend on RPI time horizon and RPI detail. For this purpose, the study varied time horizon by providing participants with RPI that is either based on the current period performance or that refers to the cumulative performance across all periods. Additionally, RPI detail was manipulated in the form that participants either received only ordinal ranking information (less detailed) or actual performance values (more detailed) on all contestants on each task. The experimental results indicate that the effort distortion of RPI is exacerbated when RPI is more detailed or when it is cumulative. Also, an interaction effect between RPI detail and time horizon is identified since effort distorting effects of more detailed RPI are greatest when RPI is cumulative. From a theoretical perspective, the authors emphasize that these

[110] Cf. Hannan et al. (2008).
[111] Cf. Eriksson et al. (2009).
[112] Cf. Hannan et al. (2013).
[113] Cf. Hannan et al. (2014)

observed effects might be driven by underlying social comparison processes and the desire for social distinction. While cumulative RPI can be more useful for social comparison purposes as it offers a more stable and meaningful measure of ability, more detailed RPI should enable participants to differentiate themselves to a greater degree from their peers. Therefore, participants might have a greater motivation to distort their effort away from an equal allocation and to concentrate on one task in which they can proof their competence.

3. Hypotheses and research questions on effects of RPI on performance and attention towards task dimensions

The objective of this study is to explore how the presence and the dimensionality of RPI affect performance and attention towards task dimensions in multidimensional task settings. In particular, the study aims to compare the effects of unidimensional RPI with that of multidimensional RPI under different compensation schemes. To address this research interest, two directed hypotheses and four two-tailed research questions are developed building on the preceding analysis of the behavioural effects of RPI.

First, the performance effects of unidimensional RPI are considered in order to establish the basis for a comparison with multidimensional RPI. From a social-psychological perspective, unidimensional RPI facilitates social comparison by which an individual can assess its overall ability to perform a multidimensional task in the absence of an objective standard. In this context, social comparison is expected to drive an individual's motivation to enhance its performance on the overall task and to reduce performance discrepancies to superior others. These motivational effects should manifest in competitive behaviour involving higher effort and also the alignment of attention towards task dimensions to the overall performance measure.[114] Therefore, unidimensional RPI is expected to positively affect performance in multidimensional task settings. Consistent with this prediction, experimental studies reviewed in chapter 2.4 provide empirical support for performance improvements induced by RPI.[115] Even though these studies relate to unidimensional task settings, they strengthen the outlined expectation since there is no theoretical indication that motivational effects of unidimensional RPI differ in multidimensional task setting. In addition, unidimensional RPI can increase learning on task-specific strategies because of its informational value as well as its potential positive effect on self-efficacy.[116] However, section 2.2.2 emphasizes that learning effects might be limited because of the multidimensionality of the task setting. Overall, the following hypothesis formally summarizes the outlined expectations:

> *H1: Overall performance is greater when individuals receive unidimensional RPI compared to when they receive no RPI.*

Next, the performance implications of providing unidimensional RPI under different compensation schemes are considered. As illustrated in chapter 2.3.1, economic theories, such as agency theory in particular, predict that performance-based compensation motivates higher performance than a flat wage which is not tied to performance.[117] In this regard, the study aims to examine the performance effects of RPI in multidimensional task settings under an individual performance-based scheme rewarding overall performance and under a flat wage contract. According to prior research, the motivational effects of financial incentives and that of RPI generally have an accumulative impact on performance.[118] Thus, given that unidimensional RPI facilitates social comparison regarding the performance measure that is relevant for compensation, performance is expected to increase when unidimensional RPI is provided together with performance-based instead of flat wage compensation. However, as outlined in section 2.3.1, it appears to be unclear whether this will be driven by an addition or a positive interaction between the motivational effects of RPI and individual performance-based compensation. Therefore, the expectations are restricted to the performance difference between the provision of unidimensional RPI under an individual performance-based scheme and a flat wage which is formally captured in the following hypothesis:

> *H2: Overall performance is greater when individuals receive unidimensional RPI and are compensated based on an individual performance-based contract compared to when they are compensated based on a flat wage contract.*

In the following, it is considered how multidimensional RPI compared to unidimensional RPI might differently affect overall performance because of its potentially deviating influence on attention towards task dimensions. Here, it seems to be preferable to develop two-tailed research questions instead of directed hypotheses. Firstly, the theoretical analysis of the behavioural effects of multidimensional RPI in chapter 2 implies conflicting influences on attention allocation which have not been studied in this context before. Secondly, the resulting overall performance cannot be clearly predicted as it is determined by the interplay between attention allocation and individual abilities.

Multidimensional RPI might have a different influence on attention allocation since it does not limit the scope of social comparison to the overall performance measure but

[114] Cf. Festinger (1954); Suls et al. (2002); Buunk and Mussweiler (2001).
[115] E.g. Blanes i Vidal and Nossol (2011); Azmat and Iriberri (2010); Tran and Zeckhauser (2012); Tafkov (2013).
[116] Cf. Kluger and Denisi (1996); Bandura (1977b); Bandura and Jourden (1991); Corcoran et al. (2011).
[117] Cf. Baiman (1982); Baiman (1990); Eisenhardt (1989).
[118] Cf. Tafkov (2013); Bonner and Sprinkle (2002); Kluger and Denisi (1996); Chung and Vickery (1976).

also facilitates peer comparisons on dimension-specific performance. As shown in section 2.1.2, this can induce a selective concentration on a task dimension in which an individual has the chance to outperform its peers and can therefore achieve self-affirmation, social distinction or satisfy its desire to win a contest.[119]

Two experimental studies, Hannan et al. (2013) and Hannan et al. (2014), find empirical support for the described behavioural effect of RPI in a multitask environment in which individuals are provided with task-specific rankings but do not receive RPI on overall performance.[120] Importantly, multidimensional RPI also compares overall performance and therefore it may only cause differences in attention allocation to the degree that it promotes a strong concentration on one of the individual task dimensions that does not reflect overall performance. However, in contrast to this motivational effect which potentially induces a distortion of attention, multidimensional RPI also implies a learning effect that may promote an appropriate allocation of attention according to individual abilities. Particularly, section 2.2.2 has outlined that multidimensional RPI can be more beneficial for learning processes on task strategies than unidimensional RPI through its informational value and its expected greater effect on self-efficacy.[121] Therefore, both an individual's motivation and learning under multidimensional RPI can cause differences in attention allocation compared to unidimensional RPI but from a theoretical perspective. However, from a theoretical perspective, it appears to be unclear how these effects ultimately manifest in behaviour. Also, so far no empirical studies provide insights on that question since Hannan et al. (2013) and Hannan et al. (2014) exclude the influence of overall rank information from their experimental setting and do not consider learning effects. Consequently, the following research question is stated:

> *RQ1a: Will attention towards task dimensions differ when individuals receive multidimensional RPI compared to when they receive unidimensional RPI?*

Elaborating on this research question, the influence of financial incentives on attention allocation differences between unidimensional and multidimensional RPI is taken into consideration. Section 2.3.2 has illustrated that a compensation contract based on overall performance implies motivational effects which might conflict with the described social-psychological incentives for a distortion of attention under multidimensional RPI. Therefore, people can face a trade-off in attention allocation to the degree that the motivations induced by the compensation contract and by multidimensional RPI are directed towards different task dimensions. The behavioural outcome of this trade-off is expected to be determined by the strength of these potentially countervailing effects which can be influenced by various individual and situational factors.[122] In this context, it can be noticed that the introduction of financial incentives can weaken attention distortion effects caused by multidimensional RPI. However, by directing attention towards overall performance, this may in turn increase the impact of learning benefits associated with the provision of multidimensional RPI. Thus, overall it cannot be predicted to which extent attention allocation differences between unidimensional and multidimensional RPI may be present when compensation is tied to individual overall performance.

This is captured in the research question below:

> *RQ1b: When individuals are compensated based on an individual performance-based contract, will attention towards task dimensions differ when they receive multidimensional RPI compared to when they receive unidimensional RPI?*

Based on the consideration of differences in attention towards task dimensions under multidimensional RPI compared to unidimensional RPI, implications for overall performance are taken into account. Importantly, there is no indication that the two forms of RPI vary in their effect on the overall level of effort expended on a task which is why the development of research questions on potential performance differences relates to differences in attention allocation. Since individuals are expected to differ in their abilities to perform various dimensions of a multidimensional task, overall performance can significantly depend on an appropriate allocation of attention. As it was emphasized before, multidimensional RPI may foster a strong concentration on a task dimension that is different from overall performance but is perceived to offer the highest chance for an individual to prove its competence and to outperform its peers. Thus, an individual is motivated to focus attention on its relative strengths within a multidimensional task which can involve positive and negative effects on overall performance. More concretely, the performance effects are determined by the marginal return of effort implied by the allocation of attention. A shift of attention toward one specific task dimension positively affects performance as long as it increases the marginal return of overall effort spend on a task compared to an alternative allocation of attention. Yet, the extent to which the marginal return can be increased by focussing on one dimension is naturally limited in multidimensional task settings. Depending on an individual's abilities, overall performance will be negatively affected when the attention devoted toward a specific dimension reaches a certain proportion at which marginal return decreases. However, it cannot be predicted to which extent individuals shift their attention based on the motivational effects of RPI. Besides, learning effects may be also a source of differences in performance resulting from the provision of the two forms of RPI. It

[119] Cf. Steele (1988); Sherman and Cohen (2006); Frey (2007); Kilduff et al. (2010).

[120] Cf. Hannan et al. (2013); Hannan et al. (2014).

[121] Cf. Kluger and Denisi (1996); Hattie and Timperley (2007); Bandura (1977b); Cervone et al. (1986).

[122] Cf. Bonner and Sprinkle (2002); Baiman (1990); Tesser (1988); Garcia et al. (2006).

was constituted before that multidimensional RPI promotes the development of a task-strategy that aligns attention towards task dimensions with an individual's abilities. Therefore, individuals are expected to be better able to maximize their marginal return to effort under multidimensional compared to unidimensional RPI. This in turn can imply possible performance improvements. In sum, the two forms of RPI should lead to different performance to the extent that their motivational and learning effects manifest in differences in the fit between attention towards task dimensions and individual abilities. On this basis, two research questions are stated which reflect the prior differentiation between possible attention allocation differences in general and in the presence of an individual performance-based compensation contract:

> *RQ2a: Will overall performance differ when individuals receive multidimensional RPI compared to when they receive unidimensional RPI?*
>
> *RQ2b: When individuals are compensated based on an individual performance-based contract, will overall performance differ when they receive multidimensional RPI compared to when they receive unidimensional RPI?*

4. Methodological approach

4.1. Laboratory experiments as a research method to study behavioural effects of RPI

In the present study, the derived hypotheses and research questions are tested within a controlled laboratory experiment. Therefore, this section briefly outlines methodical characteristics of experimental research with regard to the examination of the behavioural effects of RPI in laboratory settings.

Experimental studies in general are characterized by an active and purposeful manipulation of an independent variable in order to measure its influence on a dependent variable. By using the principle of randomization, experiments allow to control for the influence of confounding variables and hence provide the opportunity to draw causal inferences regarding the relationship between the observed variables.[123] However, compared to field experiments that take place in natural environments, laboratory experiments create artificial situations in which researchers can investigate cause-effect relations 'under pure and uncontaminated conditions'[124]. In the laboratory, the action and decision environment of experimental participants is subject to tight control and variables can be precisely and objectively measured. Accordingly, the generated results are expected to be replicable and have a higher degree of internal validity. Yet, the abstraction and simplification involved in laboratory experiments is also accompanied by a reduction of external validity in comparison to field experiments. Therefore, the possibilities to generalise results from the laboratory context to real-word settings are limited.[125]

Based on the advantage of laboratory experiments in ensuring internal validity and establishing causal inferences, they are especially useful for testing, comparing and refining theoretical predictions and assumptions.[126] Therefore, various empirical studies employ laboratory experiments in order to examine the behavioural effects of RPI.[127] The flexibility inherent in the laboratory approach allows the researchers to vary the provision of RPI in different settings, to test boundary conditions and to isolate the influence of economic and social-psychological factors.[128] In this way, it is possible to identify precise interrelations of involved variables and to gradually expand the insights on the behavioural regularities implied by RPI. Accordingly, based on previous findings, the current study conducts a laboratory experiment in order to distinctively test its hypotheses and research questions in a controlled setting and thus to provide a first understanding of the behavioural validity of theoretical predictions concerning the presence of RPI in multidimensional task environments.

As indicated before, experimental results generated in the laboratory have various desirable characteristics but researchers also face concerns regarding their external validity. Commonly, it is argued that the findings are exposed to a subject pool bias since the experimental subjects are often students. Such objections are particularly relevant for experiments that examine individual behaviour in tasks which require professional experience like for example asset trading or product pricing. Also, student samples seem to be problematic if experiments aim to describe population parameters like risk preferences or if they are interested in the influence of socio-economic factors on certain preferences or behaviour.[129] In contrast, however, the presented experimental studies on RPI as well as the current study test theoretical predictions that relate to fundamental behavioural patterns which are independent of assumptions regarding the participant pool. Therefore, it is important to note that student samples indeed lack representativeness but that this is generally not expected to be a major threat to the validity of the experimental findings on the behavioural effects of RPI.[130]

Another concern on the external validity of laboratory experiments relates to the question if they capture essential conditions that prevail in reality. Experiments cannot integrate all conditions of a real-word environment, since abstraction and simplification are necessary in order to ensure internal validity. Essentially, it depends on the purpose of the experiment to which extent the manipulations should be applied in a realistic context. Laboratory studies on RPI typically add realism by conducting 'real effort' experiments where subjects do not choose hypothetical effort

123 Cf. Sprinkle (2003), p. 289; Schnell et al. (2008); p. 212.
124 Sprinkle (2003), p. 289; cf. Kerlinger and Lee (2000).
125 Cf. Loewenstein (1999), p. 26; Schnell et al. (2008), p. 219.
126 Cf. Smith (1994); Simon (1987).
127 Cf. Chapter 2.4.
128 Cf. Sprinkle (2003), p. 289.
129 Cf. Feltovich (2011), p. 372; Falk and Fehr (2003), p. 401.
130 Cf. Falk et al. (2009), p. 55; Falk and Heckman (2009), p. 537.

levels based on a monetary function but actually work on a specific task.[131] In this context, however, researchers have to accept some loss of control since they cannot observe the participants' cost of effort. In addition, some studies examining the impact of public RPI in comparison to private RPI required experimental subjects to introduce themselves to the others at the beginning of the session.[132] Therewith the researchers removed the anonymity in the laboratory and modelled the setting closer to a workplace context in which colleagues know each other. Nevertheless, experiments cannot implement mundane realism and should only include additional context variables when they are expected to directly enhance the validity of the results since it may also reduce the superior control possibilities in the laboratory.[133]

In sum, it is important to note that results drawn from laboratory experiments can be usefully complemented with findings from field experiments or quasi-experiments.

Building on clean evidence generated in the laboratory, field data can provide representative insights since they capture individual's behaviour in real situations and typically include larger samples.[134] Therefore, research progress regarding the behavioural effects of RPI is driven by both laboratory and field studies.

4.2. Specification of the experimental design

4.2.1. Real effort task

In accordance with previous laboratory experiments examining the behavioural effects of RPI, this study employs a 'real effort' experiment in order to test its hypotheses and research questions. The current chapter hence starts with a description of the applied 'real effort'-task which is then used a basis to explain the experimental manipulations.

Participants of the experiment are required to perform a multidimensional task on an individual computer terminal for six consecutive rounds of 360 seconds each.[135] The task consists of solving multiplication problems with different difficulty levels. In particular, participants are provided with task blocs that include six multiplication problems which are ordered according to increasing difficulty. All problems of a bloc are shown on one screen but participants can only solve them in the given order. Therefore, participants have to start each task bloc with the problem of the lowest difficulty level and can only proceed to the next level when they solve it correctly. After entering a solution to a problem, participants have to click on the corresponding 'OK'-button so that the program can check whether the answer is correct or not. A warning message appears in case participants enter wrong answers and then they have the opportunity to retry solving the problem. Importantly, participants do not have to deal with all difficulty levels included in a bloc but can also change to the next task bloc after solving the first problem of the lowest difficulty level. However, participants receive points for solving the multiplication problems and the higher the difficulty level, the more points can be earned on a problem. Therefore, the points per task bloc rise progressively with the difficulty level. Exemplary, a task bloc with two correctly solved problems accounts for 3 points while a bloc with four problems accounts for 16 points. Once participants click on the corresponding button to change to the next bloc, they are prohibited from returning to the previous one. Per round, a maximum of 75 task blocs can be displayed which is yet far beyond the number of blocs that participants can solve in 360 seconds even if they only work on problems with the lowest difficulty level. The task blocs are presented in the same order to all participants and have to be solved without any outside aid which also includes writing materials.

The total number of points earned by participants reflects their overall task performance and is determined by their performance in two task dimensions – the number of task blocs solved as well as the average number of points earned per bloc. Participants can individually decide at which difficulty level they change from one task bloc to another and therefore they have the opportunity to adopt various task strategies. Notably, there is no dominant strategy since participants' overall performance depend on the extent to which they align their attention towards the two task dimensions with their abilities in solving multiplication problems. For example, a certain participant may make the best out of his abilities by solving many task blocs with a rather low number of points per bloc. In contrast to such a 'quantity'-strategy, another participant can be able to maximize his overall performance by following a 'value'-strategy which implies that he solves fewer task blocs but concentrates on achieving a higher number of points per bloc.

Following a math task applied by Hannan et al. (2013), the differentiation between the difficulty levels of the multiplication problems is based on two criteria: (1) number of digits in the two multiplicands (e.g. multiplying a one-digit and a two-digit number compared to multiplying two, two-digit numbers) and (2) frequency of carry-overs required to solve the problem.[136] Various pre-tests have been conducted to specify the two criteria in a manner that the time required for solving a problem increases considerably with the difficulty level. Additionally, the increase in points that can be earned per difficulty level is balanced so that the marginal return of solving an additional problem differs considerably with the participants' abilities. Specifically, the time that it takes a participant to solve a problem determines the difficulty level until which there is an increase in points earned per second of time that is spend on a task bloc. Overall, this ensures that the task design is appropriate for testing the underlying theory.

[131] Cf. Falk and Fehr (2003), p. 404; v. Dijk et al. (2001).

[132] Cf. Tafkov (2013); Hannan et al. (2013).

[133] Cf. Libby et al. (2002), p. 797.

[134] Cf. Falk et al. (2009), pp. 5-6; Falk and Ichino (2006); Azmat and Iriberri (2010).

[135] The experiment was programmed using z-Tree software, cf. Fischbacher (2007).

[136] Cf. Hannan et al. (2013), p. 559-560. See Online-Appendix 1 for an overview on the specification of the difficulty levels, including the number of points that participants could earn per difficulty level.

4.2.2. Experimental manipulations

The conducted experiment is characterized by a 3 x 2 between-subjects design that implies the manipulation of two factors which represent the primary independent variables of interest. The first factor constitutes the provision of RPI and is varied a three levels: No RPI, Unidimensional RPI and Multidimensional RPI. The second factor is the compensation contract which is manipulated at two levels: Flat wage contract and individual performance-based contract. Therefore, the experiment involves six treatments which are applied between-subjects so that each of the participants is exposed to one of the treatments.

The use of a between-subject design has the advantage over a within-subjects design that carryover-effects between treatments are avoided which can otherwise threaten the internal validity of the results. Additionally, a between-subject design reduces 'demand effect' concerns which relate to the risk that participants form assumptions over an experiment's purpose and change their behaviour accordingly. Therefore, previous experiments examining the behavioural effects of RPI systematically rely on between-subjects design. However, it should be also noted that within-subjects designs in comparison have the upside that fewer participants are required to generate statistical powerful results since it is controlled for subject-specific noise.[137]

In all treatments, participants receive individual performance information at the end of each round showing them their overall score as well as the number of task blocs solved and the average number of points earned per task bloc. The latter only takes into account task blocs that are completed, meaning that participants have clicked on the corresponding button to change to the next bloc. Therewith, it is ensured that the performance measure is not distorted by unfinished task blocs at the end of the rounds. The individual performance information is provided with regard to the current period performance and the cumulative performance across all periods. During all rounds, participants can keep track of their current period performance which is displayed on the top of the task screen.

In the 'No RPI' treatment, no additional performance information is presented to the participants. In the treatments that involve the provision of RPI, participants are assigned to groups of five and receive public ranking information that compare their cumulative performance to that of the other group members at the end of each round. For that purpose, participants have to choose one out of eight given fictional nicknames at the beginning of the experiment which they keep over all rounds.[138]

Since anonymity has to be preserved in order to comply with the guidelines of the laboratory in which the experiment is conducted, participants do not know how the chosen names match to the actual persons in the laboratory. In general, the presented RPI only informs participants about performance ranks but they do not learn about the specific performance levels of the other group members. In this way, RPI cannot be used for wage comparison purposes in the treatment that applies an individual performance-based contract.

Participants in the 'Unidimensional RPI' treatment receive only one ranking comparing the cumulated number of points of the group members after every round. In the 'Multidimensional RPI' treatment, two additional rankings are provided that involve performance comparisons related to the number of task blocs solved and the average number of points earned per task bloc. In both treatments, participants with the highest rank after a certain round see a picture of a golden trophy on their computer screen next to the corresponding ranking. At the end of the experiment, congratulations to the final winners of the rankings are displayed publicly in addition to the privately shown trophies. Through the combination of RPI with such symbolic rewards, the setting is modelled closer to a 'gamification' context in which employees receive virtual batches that demonstrate their relative success in a certain task or task dimension.[139]

As indicated before, the manipulations related to the provision of RPI are crossed by the variation of the compensation contract at two levels. In the 'Flat wage' treatment, participants earn a fixed amount of 1 Euro for each of the six rounds independent of their performance on the task. In contrast, the compensation in the 'Individual performance-based' treatment is determined by the total number of points that participants have achieved at the end of the experiment. Specifically, they receive 0.2 Euro for each point which implies that they can earn a maximum of 1.10 Euro for a completely solved task bloc which accounts for 55 points.

4.3. Experimental procedures and participants

The experiment was conducted in the experimental laboratory MELESSA[140] at the University of Munich. It involved six separate experimental sessions with each representing one of the described treatments.

The sessions were divided into two independent parts – the first part was designed to elicit the participants' risk preferences and the second part represented the actual real effort experiment. Upon arrival, participants were told that they would go through two independent experiments and then received instructions on the first part which were read aloud by the experimenter.[141] Participants were not informed what the second part would involve. The measurement of the participants' risk preferences in advance to the real effort experiment is necessary in order to control for the influence of risk preferences on attention towards task dimensions. In

[137]Cf. Libby et al. (2002), p. 804; Schnell et al. (2008), p. 210; Field and Hole (2002), p. 75.

[138]The given nicknames include gender-neutral common surnames which did not allow for conclusions on personal characteristics. The members of a group are required to choose different names.

[139]Cf. Deterding et al. (2011); Reeves and Read (2013).

[140]MELESSA abbreviates 'Munich Experimental Laboratory for Economic and Social Sciences'.

[141]See Online-Appendix 2 for the instructions on the first part of the experiment.

particular, the risk involved in solving multiplication problems increases with the difficulty level of the problems since participants get more uncertain about whether they will be able to solve the problem correctly and about how long it will take. Therefore, risk aversion might be a determinant of attention allocation when compensation is tied to individual performance.[142] The experimental design that is applied to elicit participants' risk preferences is similar to one developed by Holt and Laury (2002) and consists of simple lottery choices.[143] Participants were required to make ten choices between paired lotteries and were informed that one choice will become relevant for payment by which the incentive compatibility of each decision is ensured. The information on the relevant choice as well as the associated payment were revealed at the end of the experimental session.

Upon completion of the first part, the instructions for the actual real effort experiment were handed out and read aloud.[144] Participants had the chance to ask questions and were then required to take a pre-experiment quiz. So as to ensure that participants understand the experiment, they could not proceed until they scored 100 percent on the quiz.[145] After all participants passed the quiz, they were given the opportunity to practice the task in a training round in which no manipulations were applied. The instructions emphasized that participants should use the training round in order to develop a task strategy that fits their abilities. Following to the training round, participants could again ask questions in case anything remained unclear to them. Next, participants in the treatment conditions that involved the provision of RPI had to choose a fictional name which they kept for the rest of the experiment. Then, participants performed the real effort task for six consecutive rounds. Afterwards, they were asked to complete a post-experiment questionnaire that included questions concerning controls and demographic data.[146] Finally, participants privately received their payment from the experiment in cash. On average, participants earned 12.7 Euro for approximately one hour of participation, including the payment from the lottery of averagely 2.5 Euro as well as the show-up fee of 4 Euro. Notably, the average earnings did not differ significantly between the two compensation treatments (t=0.54, p=0.58, two-tailed).

The experiment was conducted with 150 participants so that each session employed 25 participants. All participants were students that had been recruited from the participant pool of the experimental laboratory MELESSA through an anonymous invitation procedure using ORSEE software[147]. Of the participants, 86 were male (57.3 percent) and 64 were female (42.7 percent). Overall, participants were similar on related attributes like age, education and mathematical background. The mean age was 22.9 (s.d.=3.66) and the participants were primarily undergraduate students (74.7 percent)[148]. Furthermore, most of them majored in a math-related field of study (80.6% percent)[149]. With regard to their risk preferences, the results of the first experiment indicate that the participants were overall risk averse. On average, their ten lottery choices involved 5.99 (s.d.=1.64) save choices.[150]

Additionally, the participants' preference for competition is measured in the post-experimental questionnaire by a question adapted from Griffin-Pierson (1990).[151] Using a seven-point Likert scale, participants had to assess the degree to which they perform better when they are competing against someone rather than when they are the only one striving for a goal (1=not at all, 7=to a great degree). The participants' average assessment equaled 4.09 (s.d.=1.99). Furthermore, the questionnaire included control questions which required participants to indicate on the same scale the degree to which they were interested in the task (mean=4.23, s.d.=2.06) and to which they knew the task from prior experiments (mean=1.29, s.d.=0.95).

Across the treatments, there was no significant variation in the presented personal variables except for some differences in gender, age, mathematical background and risk aversion.[152] Therefore, the multivariate analysis of the experimental results controls for the influence of these individual characteristics on performance as well as on attention towards task dimensions.

5. Results

5.1. Tests of hypotheses and research questions

5.1.1. Bivariate analyses

The two primary dependent variables of the analysis of the experimental results are performance and attention towards task dimensions. The performance of the participants is measured by the total number of points that they earned through solving multiplication problems over six rounds. The

[142] Agency theory emphasizes that risk-averse individuals dislike the income risk that is associated with performance-based compensation (cf. Baiman (1990), Prendergast (1999)). In the applied set- ting, participants can reduce their income risk by concentrating on multiplication problems with low difficulty level.

[143] Cf. Holt and Laury (2002).

[144] See Online-Appendix 3 for the instructions on the real effort experiment.

[145] See Online-Appendix 4 for the questions of the pre-experiment quiz.

[146] See Online-Appendix 5 for the questions of the post-experimental questionnaire.

[147] Cf. Greiner (2004).

[148] The rest of the participants either had already completed a bachelor's degree (16.0 percent) or a master's degree/diploma (9.3 percent).

[149] The math-related fields of study include Business & Economics (55.3 percent), Natural Science & Engineering (17.3 percent) as well as Formal Sciences (8.0 percent).

[150] Cf. Holt and Laury (2002), p. 1649, 5 save choices indicate 'slight risk aversion', 6 save choices indi- cate 'risk aversion'.

[151] Cf. Griffin-Pierson (1990).

[152] 1) Gender: 40% male participants under Multidimensional RPI/Flat Wage, 64% under Unidimen- sional/Flat Wage, No RPI/Flat Wage, No RPI/Perf.-based (t=1.71, p=0.09, two-tailed). 2) Age: Mean age of 21.76 under No RPI/Perf.-based, 23.88 under Multidimensional RPI/Perf.-based (t=1.98, p=0.06, two-tailed). 3) Math Background: 68% of participants major in math-related field of study under Unidimen- sional RPI/Flat Wage, 92% under No RPI/Perf.-based (t=2.18, p=0.04, two-tailed). 4) Risk Preferences: Mean number of save choices is 5.68 under No RPI/Flat Wage, 6.64 under Unidimensional RPI/Flat Wage (t=2.10, p=0.04, two-tailed).

measure of attention towards task dimensions is defined by the participants' average difficulty level per task bloc which underlies their overall performance.[153] This variable reflects the participants' allocation of attention between the 'value' dimension and the 'quantity' dimension of the task since it indicates at which difficulty level participants have changed from one task bloc to another.[154]

Table 1 and Table 2 report summary statistics for performance as well as attention towards task dimensions and provide bivariate comparisons between treatments. In Panel A of each of the Tables, treatment data are pooled for the three levels of RPI. Panel B presents a comprehensive overview on all six treatments including the effects of the variation of the compensation contract. Figure 1 and Figure 2 each show a graphical summary of the mean values presented in Panel B of Table 1 and Table 2. Since both dependent variables do not follow a normal distribution[155], non-parametric tests are used to statistically compare treatments. First, p-values are reported for a two-tailed Mann-Whitney pairwise test which examines whether two independent samples are drawn from populations with the same distribution.[156] In order to allow for a more comprehensive analysis, the results of a Pearson Chi-Squared test on the equality of medians[157] are provided in addition because the data samples on the dependent variables both include various (extreme) outlying observations[158]. Overall, each participant is treated as an independent observation which also applies to further statistical tests that are used in the course of the analysis.[159]

In hypothesis H1, it is predicted that participants perform better under the provision of unidimensional RPI compared to when they do not receive RPI. The data in Panel A of Table 1 support this prediction since they show that performance is significantly higher under unidimensional RPI (p=0.06 / 0.05)[160]. Notably, the standard deviation of the performance of participants provided with unidimensional RPI is also considerably higher which reflects a greater spread of the data.[161] Panel B of Table 1 shows that this spread is especially driven by the treatment involving performance-based compensation since the related standard deviation of performance clearly exceeds the one that can be observed under a flat wage contract. Furthermore, in accordance with the prediction of hypothesis H2, mean values indicate a performance increase when participants receive performance-based instead of flat wage compensation in addition to unidimensional RPI. As illustrated in Figure 1, this performance effect seems to be additive since the induced performance difference between the two compensation contracts is similar when participants are not presented with RPI. However, in contrast to the performance effect suggested by the mean values, the related medians imply a slightly negative effect of performance-based compensation. This difference between mean and median values can be explained by the strongly right-skewed distribution of performance under a performance-based contract[162] which includes two extreme outliers[163]. Consequently, the p-values obtained from the comparison between performance-based and flat wage compensation under unidimensional RPI reveal that there is no significant performance difference (p=0.82 / 0.40). Therefore, in sum, the results do not support the prediction of hypothesis H2.

The research questions RQ2a and RQ2b address the performance effect of multidimensional RPI compared to that of unidimensional RPI. First, with regard to RQ2a, Panel A of Table 1 suggests that participants perform worse when they receive multidimensional instead of unidimensional RPI. Nevertheless, the reported p-values narrowly fail to confirm a significant performance difference between the two forms of RPI (p=0.19 / 0.11). Yet, participants provided with multidimensional RPI also do not perform significantly better than those in the 'No RPI' treatments although this is suggested by mean and median values (p=0.79 / 0.69)[164]. This contrasts the significant performance effect of unidimensional RPI which was presented before. Further insights on the influence of multidimensional RPI on performance can be drawn from Panel B of Table 1. Importantly, summary statistics indicate that the performance decrease under multidimensional compared to unidimensional RPI is more pronounced when participants are compensated with a performance-based contract. This interaction between the dimensionality of RPI and the applied financial incentives is graphed in Figure 1. In his context and with regard to RQ2b, the Pearson Chi-Squared test

[153]In the following, 'average difficulty level', 'difficulty level' and 'attention towards task dimensions' is used synonymously.

[154]The average difficulty level is calculated based on the averages per round. The last task bloc of a certain round is excluded from the calculation of the average difficulty level per round in order to avoid distortion because of unfinished task blocs. When only one task bloc was solved in a period (relates to 22 out of 900 observations), the difficulty level of this bloc is included.

[155]Normality tests on performance / attention towards task dimensions: 1) Shapiro-Wilk-Test: z=7.14 / 5.29, p=0.00 / 0.00; 2) Skewness/Kurtosis tests: adj. $\chi^2(2)$=73.47 / 23.76, p=0.00 / 0.00, Skewness=2.48 / 1.16, Kurtosis=12.55 / 4.13.

[156]Cf. Mann and Whitney (1947); Acock (2006), p. 144.

[157]Cf. Brown and Mood (1951); Acock (2006), p. 145; Vorberg and Blankenberger (1999), pp. 158-160

[158]Outliers are defined as observations which are at least 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile of the distribution, cf. Hoaglin et al. (1986). Extreme outliers are at least 3.0 IQRs below the first quartile or above the third quartile. See Online-Appendix 6 and Online-Appendix 7 for boxplots on the dependent variables performance and difficulty level by treat- ment which illustrate the (extreme) outlying observations.

[159]This corresponds to an approach used by prior experimental studies (e.g. Hannan et al. (2008), Bracha and Fershtman (2013)). However, it is controversial whether individual-level observations within a ranking group are independent from each other.

[160]Order of p-values: Mann-Whitney test / Pearson Chi-Squared test.

[161]Levene's test indicate that there is overall no significant variance heterogeneity between treat- ments: F=0.86, p=0.51 (cf. Levene (1960), results are robust under nonnormality).

[162]Skewness/Kurtosis tests on performance under unidimensional RPI and performance-based contract / flat wage contract: adj. $\chi^2(2)$=22.57 / 3.86, p=0.00 / 0.15, Skewness=2.50 / 0.63. See Online-Appendix 8 and Online-Appendix 9 for histograms on the dependent variables performance and difficulty level by treatment.

[163]The extreme outlying observations relate to 2 participants with total points of 827 (IQR: 3.28) and 1312 (IQR: 6.79).

[164]P-values are not included in Table 1.

Table 1: Dependent Variable Performance – Summary Statistics and Test of Hypotheses and Research Questions

mwt = Mann-Whitney test.
pct = Pearson Chi-squared test on the equality of medians.
P-values compare treatment groups and treatments according to the formulated hypotheses and research questions.
Treatment groups: Unidimensional RPI compared No RPI; Multidimensional RPI compared to Unidimensional RPI.
Treatments: No RPI compared within treatment group; Unidimensional RPI/Perf.-based compared within treatment group to Flat Wage, Unidimensional RPI/Flat Wage compared between treatment groups towards No RPI/Flat Wage.
Multidimensional RPI compared between treatment groups towards Unidimensional RPI.
* Significant at 10%; ** Significant at 5%; *** Significant at 1% level.

RPI		**None**		**Unidimensional**		**Multidimensional**	
Panel A	Mean	285,74		347,74		312,6	
	Median	255		327,5		275,5	
	s.d.	128,3997		209,5149		188,3382	
	mwt	-		0.06* [H1]		0.19 [RQ2a]	
	pct	-		0.05** [H1]		0.11 [RQ2a]	
Compensation Contract		Flat Wage	Perf.-based	Flat Wage	Perf.-based	Flat Wage	Perf.-based
Panel B	Mean	271,72	299,76	332,12	363,36	323,04	302,16
	Median	249	300	331	321	315	234
	s.d.	128,0487	129,8227	164,2331	249,2813	147,8551	224,3486
	Range	144-772	107-620	23-744	87-1312	134-762	88-1217
	n	25	25	25	25	25	25
	mwt	-	0,35	0.05**	0.82 [H2]	0.67 [RQ2b]	0.11 [RQ2b]
	pct	-	0,16	0.01***	0.40 [H2]	0.4 [RQ2b]	0.09* [RQ2b]

Table 2: Dependent Variable Difficulty Level – Summary Statistics and Test of Hypotheses and Research Questions

mwt = Mann-Whitney test.
pct = Pearson Chi-squared test on the equality of medians.

RPI		**None**		**Unidimensional**		**Multidimensional**	
Panel A	Mean	2,77		2,88		2,75	
	Median	2,69		2,51		2,58	
	s.d.	0,97		1,16		1,05	
	mwt	-		0,95		0,72	
	pct	-		0,69		1	
Compensation Contract		Flat Wage	Performance-based	Flat Wage	Performance-based	Flat Wage	Performance-based
Panel B	Mean	2,53448	3	2,58	3,18	2,69	2,81
	Median	2,26	2,75	2,31	2,75	2,64	2,51
	s.d.	0,8371085	1,04	0,98	1,28	1,14	0,98
	Range	1.01 - 4.75	1.61 - 5.67	1.23 - 5.58	1.86 - 6.00	1.13 - 6.00	1.57 - 6.00
	mwt	-	0,11	0,93	0.10*	0,81	0,36
	pct	-	0,16	0,78	0,16	0,4	0,16

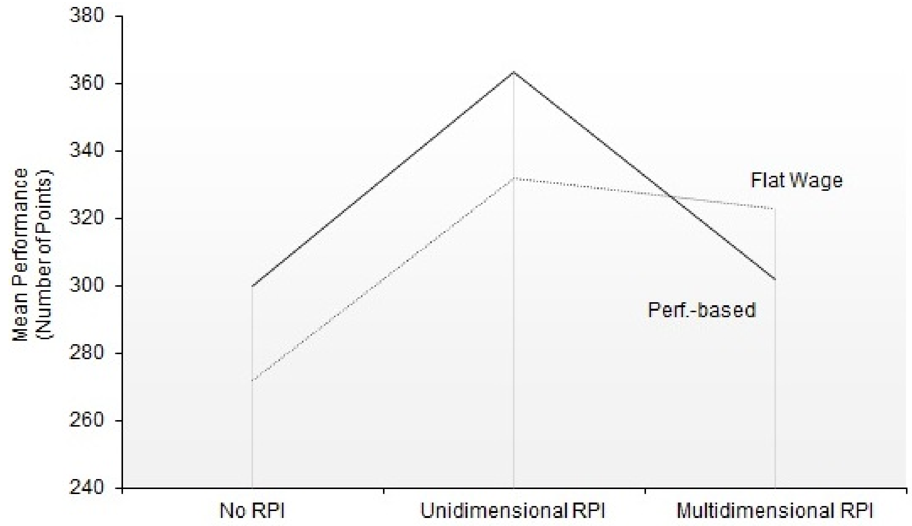

Figure 1: Effect of RPI and Financial Incentives on Mean Performance

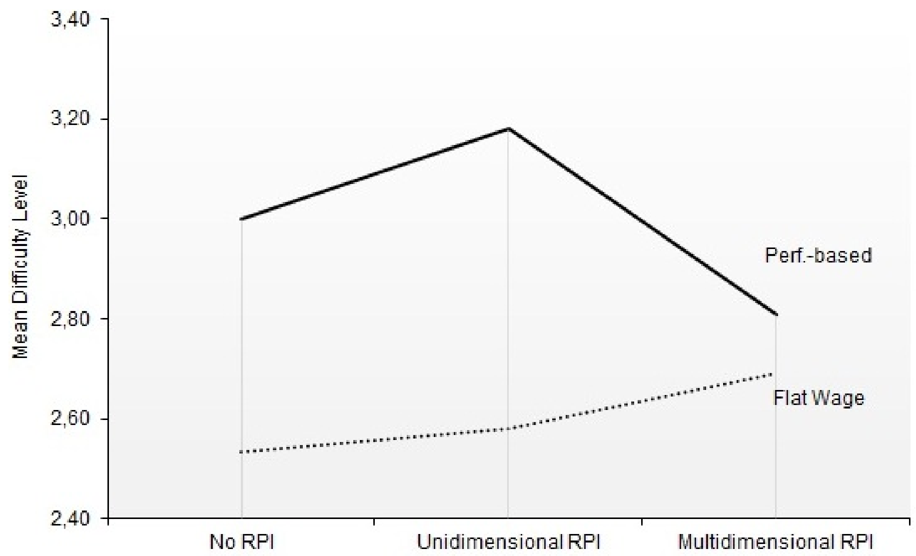

Figure 2: Effect of RPI and Financial Incentives on Mean Difficulty Level

suggests that median performance is significantly different when multidimensional RPI instead of unidimensional RPI is used in conjunction with performance-based compensation (p=0.11 / 0.09).[165] Additionally, it should be noted that the standard deviation of performance is also remarkably higher under a performance-based contract compared to a flat wage contract which is similar to the characteristics of the data on performance under unidimensional RPI.

The examined research questions which are concerned with the effects of the dimensionality of RPI on performance are derived from the research questions RQ1a and RQ1b that deal with possibly varying influences of unidimensional and multidimensional RPI on attention towards task dimensions. Considering RQ1a, the data in Panel A of Table 2 reveal that the mean difficulty level at which participants changed from one bloc to another is lower under multidimensional compared to unidimensional RPI. However, it is also shown that the median difficulty level is slightly higher under multidimensional RPI. Overall, there is no significant difference in attention towards task dimensions between the two forms of RPI (p=0.72 / 1.00). Also, a comparison between the treatments that involve the provision of multidimensional RPI and the 'No RPI' treatments does not imply any significant variation (p=0.69 / 0.69)[166].

The puzzling results derived from the pooled data are driven by differences in the impact of the two forms of RPI

[165]The Mann-Whitney test do not indicate significance because of an upward outlier under multidimensional RPI and performance-based compensation. The related participant reached a total number of points of 1217 (IQR: 6.38).

[166]P-values are not included in Table 2.

on attention towards task dimensions under performance-based and flat wage compensation. As reported in Panel B of Table 2, the application of a performance-based contract is associated with an increase in mean and median difficulty level when participants receive unidimensional RPI or no RPI. The Mann-Whitney test indicates that this difference in attention allocation between the compensation contracts is significant for unidimensional RPI (p=0.10 / 0.16).[167] With regard to multidimensional RPI, the mean difficulty levels suggest a slight increase when performance-based compensation is applied but the median values show a reduction. The indication of a positive effect is caused by an extreme upward outlier[168] under performance-based pay and differences in the skewness of the distributions.[169] A comparison between the effects of the two forms of RPI under different financial incentives reveal that multidimensional compared to unidimensional RPI induces a higher difficulty level under a flat wage compensation but a lower difficulty level in the presence of a performance-based contract. Figure 2 graphically summarizes the described interaction between the dimensionality of RPI and the used compensation contract. However, concerning RQ1b, no significant difference can be found in attention towards task dimension when participants receive multidimensional compared to unidimensional RPI under performance-based compensation (p=0.36 / 0.16). Likewise, there is no significant difference when a flat wage contract is applied (p=0.81 / 0.40).

So far, it can be summarized that unidimensional RPI significantly increases participant's performance in the examined multidimensional task setting even though this effect is not magnified under performance-based compensation which contradicts the formulated prediction. Importantly, the results of the bivariate analysis indicate that multidimensional RPI has no significant positive effect on performance but that it rather induces a performance decrease compared to unidimensional RPI in the presence of a performance-based contract. The negative performance effect is accompanied by a lower difficulty level at which participants changed from one task bloc another. However, this difference in attention towards task dimensions between the two forms of RPI under performance-based compensation appears not to be significant. Overall, it can be also noted that both the provision of unidimensional and multidimensional RPI results in a considerably higher spread of performance as compared to when participants receive no RPI.

5.1.2. Multivariate analyses

Based on the results of the bivariate analyses, this section aims to provide a more differentiated understanding of the effects of the dimensionality of RPI in the presence of financial incentives. For this purpose, multiple linear regressions are conducted which control for further determinants of the dependent variables and consider interaction effects. In particular, three models are specified which are tested with an OLS regression as well as a quantile regression approach. The latter is defined for the 50th percentile and involves median regressions. Notably, it is included in the analysis because of its non-parametric character and its robustness to outlying observations.[170]

In Table 3, results are reported for the dependent variable performance. The findings of the median regression analysis are only shown for model (3) since its results for model (1) and model (2) do not considerably differ from the presented OLS estimations.[171] In order to ensure the robustness of the OLS regression results, heteroskedasticity-consistent standard error estimators are implemented.[172] In this context, it should be also noted that post-estimation analyses indicate no threat of multicollinearity to the results.[173]

Model (1) consists of dummy variables indicating the provision of unidimensional RPI, multidimensional RPI as well as performance-based compensation. The results imply that performance is significantly higher under unidimensional RPI (p=0.077) while the positive performance difference between multidimensional RPI and no RPI is not statistically significant (p=0.408). Also, the influence of performance-based compensation appears to be positive but the coefficient is insignificant (p=0.663). However, overall, the explanatory power of model (1) is low and the null hypothesis cannot be rejected that all of the model coefficients are zero (Prob>F=0.354, R^2=0.021).

In model (2), personal variables are added so as to control for differences in participants' characteristics between treatments. Specifically, the participants' age and risk aversion are included as well dummy variables that divide for gender and indicate whether participants major in a math-related field of study. Additionally, model (2) controls for participants' preferences for competition because prior literature suggest them to be determinant of performance in competitive task settings.[174] Importantly, the results reveal that the effects of the dimensionality of RPI and of the applied compensa-

[167]The Pearson Chi-Squared test does not indicate significance since the difference is strongly driven by two observations which relate to a participants with an average difficulty level of 5.7 and 6.0.

[168]The related participant reached an average difficulty level of 6 (IQR: 6.81).

[169]Skewness/Kurtosis tests on attention towards task dimensions under multidimensional RPI and performance-based contract / flat wage contract: adj. $\chi^2(2)$=7.42 / 13.48, p=0.03 / 0.00, Skew- ness=1.66 / 1.08. See Online-Appendix 8 and Online-Appendix 9 for histograms on the dependent variables performance and difficulty level by treatment.

[170]Cf. Koenker and Bassett (1978); Koenker (2005).

[171]A comprehensive overview of the results of the median regressions for model (1) - (3) can be found in Online-Appendix 10.

[172]Breusch-Pagan test (cf. Breusch and Pagan (1979)) and White's test (cf. White (1980)) are applied in order to check for the homoscedasticity of residuals. The Breusch-Pagan test (normality assump- tion removed) indicates heteroskedasticity for model (2) and model (3) ($\chi^2(1)$=5.51, p=0.019; $\chi^2(1)$=8.87, p=0.003).

[173]Estimated variance inflation factors (VIF) are constantly below the value 10 and thus indicate that there is no problematic degree of multicollinearity (cf. Marquaridt (1970)). VIF ranges are as follows: model (1): 1.00-1.33; model (2): 1.02-1.43; model (3): VIF=1.11-3.39.

[174]E.g. prior studies are concerned with gender differences in preferences

Table 3: Dependent Variable Difficulty Level – Regression Results

Model (1) reports treatment group effects only. Model (2) additionally includes the control variables Gender (dummy variable, 1='male', 2='female'), Age (mean-centered), Math-related Field of Study (dummy variable, 1='Math-related Field of Study', 2='No Math-related Field of Study), Risk Aversion (range 0 - 10 where 0 indicates 'highly risk loving' and 10 indicates 'highly risk averse'), Preference for Competition (participants' responses on a seven-point Likert scale to the statement 'I perform better when I compete against someone rather than when I am the only one striving for a goal', 1='not at all', 7='to a great degree'). Model (3) adds interaction terms between the treatment groups for RPI and for compensation scheme as well as interaction terms between the treatment groups for RPI and a dummy variable for participants' position in the final ranking on overall performance (1='Rank 4 or 5', 2='Rank 1, 2 or 3'). Heteroskedasticity-consistent standard error estimators are implemented in the OLS regression.
* Significant at 10%; ** Significant at 5%; *** Significant at 1% level.

Regressors (DV: Performance)	**Model (1) OLS Reg.**	**Model (2) OLS Reg.**	**Model (3) OLS Reg.**	**Model (4) OLS Reg.**
	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value
Unidimensional RPI	62 (34.782) 0.077*	71.173 (39.136) 0.071*	159.7 (45.774) 0.001***	149.72 (27.287) 0.000***
Multidimensional RPI	26.86 (32.358) 0.4078	14.199 (32.143) 0.6594	128.8 (41.38) 0.002***	99.212 (26.656) 0.000***
Perf.-based Compensation	12.8 (29.27) 0.6625	20.139 (27.806) 0.4701	43.008 (35.829) 0.2321	74.395 (24.833) 0.003***
Gender		2.7174 (27.933) 0.9226	10.682 (23.853) 0.655	-6.598 (15.269) 0.6664
Age		10.473 (4.0629) 0.011**	9.3839 (3.8275) 0.015**	6.3 (2.0672) 0.003***
Math-related Field of Study		63.401 (40.659) 0.1212	24.004 (34.228) 0.4843	45.385 (19.366) 0.021**
Risk aversion		-14.34 (10.517) 0.1748	-15.73 (8.587) 0.069*	-10.25 (4.3635) 0.020**
Preference for Competition		15.696 (7.7417) 0.044**	5.6771 (7.0996) 0.4253	2.1902 (3.7655) 0.5618
Unidimensional RPI x Perf.-based Compensation			-12.95 (56.512) 0.819	-69.86 (35.073) 0.048**

(Continued)

for competition that impact performance outcomes, cf. Croson and Gneezy (2009); Niederle and Vesterlund (2007).

Table 3—continued

Regressors (DV: Performance)	**Model (1) OLS Reg.**	**Model (2) OLS Reg.**	**Model (3) OLS Reg.**	**Model (4) OLS Reg.**
Multidimensional RPI x Perf.-based Compensation			-67.1 (56.761) 0.2392	-91.62 (35.112) 0.010***
Unidimensional RPI x Low Rank			-208.5 (46.048) 0.000***	-132.3 (25.535) 0.000***
Multidimensional RPI x Low Rank			-197.2 (41.653) 0.000***	-169 (26.248) 0.000***
Constant	279.34 (23.245) 0.000***	245.85 (78.173) 0.002***	310.64 (72.296) 0.000***	253.21 (34.588) 0.000***
N	150	150	150	150
Prob>F	0.3538	0.0098	0	-
R^2	0.0214	0.1108	0.3025	-
Pseudo R^2	-	-	-	0.2145

tion contract remain largely unchanged when personal variables are considered. Yet, it is also shown that participants' age as well as their preference for competition both significantly positively affect performance (p=0.011, p=0.044). Furthermore, model (2) in general is statistical significant and has thus a higher explanatory power than model (1) (Prob>F=0.010, R^2=0.111).

Model (3) introduces interaction terms which account for differences in the performance effects of the dimensionality of RPI under performance-based and flat wage compensation. Also, it contains interaction variables that distinguish between the influences of the two forms of RPI on high ranked and low ranked participants.[175] In this way, the model addresses findings from the bivariate analyses which indicate that performance effects of RPI differ with regard to the applied compensation contract as well as that the standard deviation of performance is comparably higher in the presence of RPI.

Overall, the inclusion of the interaction terms substantially increases the predictive power of the model (Prob>F =0.000, R^2=0.303). The results are reported for an OLS as well as a median regression. Importantly, both statistical analyses indicate that performance significantly rises under unidimensional RPI (p=0.001 / 0.000)[176] as well as under multidimensional RPI (p=0.002 / 0.000). The median regression finds moreover a significant positive effect of performance-based compensation (p=0.003) while the related OLS estimates are still insignificant (p=0.232). By contrast, however, the interactions between performance-based compensation and unidimensional as well as multidimensional RPI are both negative (p=0.819 / 0.048; p=0.239 / 0.010)[177]. The results are again significant for the median regression but not for the OLS regression which may be explained by the already mentioned extreme upward outliers that are present in the treatments that involve RPI and performance-based compensation. Notably, the size of the coefficient of the interaction term is larger for multidimensional RPI than for unidimensional RPI. In this context, the significant negative interaction between multidimensional RPI and performance-based pay in the median regression has almost the same size as the positive performance effect indicated by the dummy variable for multidimensional RPI. Therefore, the overall effect of the application of multidimensional RPI under performance-based compensation is nearly zero. As opposed to that, OLS and median regression results both imply that unidimensional RPI overall impacts performance considerably positively when performance-based compensation is applied.

Further on, model (3) shows significant negative interactions between the two forms of RPI and the overall rank-

[175] High ranked participants are characterized by a final rank of 1, 2 or 3 on overall performance. Low ranked participants are characterized by a final rank of 4 or 5 on overall performance.

[176] Order of p-values: OLS regression / Median regression.

[177] The p-values that refer to the effect of unidimensional RPI are mentioned first.

ing position of the participants (p=0.000 / 0.000; p=0.000 / 0.000). Specifically, in the case of multidimensional RPI, the size of the coefficients of the OLS as well as the median regression indicate that the performance of low ranked participants overall decreases as compared to when they receive no RPI. For unidimensional RPI, the same can be concluded with regard to the estimates of the OLS regression. Also, the negative interaction effect appears to be greater for unidimensional than for multidimensional RPI considering the OLS results. Yet, the opposite holds for the findings of the median regression. Additional analyses disclose that these discrepancies between the statistical analyses are caused by a downward outlier[178] included in the observations on performance under unidimensional RPI which relates to a low ranked participant. Lastly, with respect to the influence of personal variables on performance, model (3) confirms the positive influence of participants' age (p=0.015 / 0.003) but also predicts a significant negative effect of risk aversion (p=0.069 / 0.020). Furthermore, the positive impact of participants' preference for competition is no longer significant (p=0.425 / 0.562) while the median regression indicates that majoring in a math-related field of study significantly increases performance (p=0.484 / 0.021).

Overall, post-estimation analyses applied to the presented models reveal that the performance effects of unidimensional and multidimensional RPI under flat wage compensation do not significantly differ from each other.[179] Similarly, in model (3), no significant differences are found between the interaction effects of the two forms of RPI with the applied compensation scheme[180] and the overall ranking position of the participants[181].

Table 4 reports regression results for the dependent variable attention towards task dimensions. The presented models correspond to the models applied in the regression analysis of the dependent variable performance. The estimates of the median regression are not included in Table 4 since they do not involve additional insights.[182] Like in the prior analysis on performance, heteroskedasticity-consistent standard error estimators are used with regard to the robustness of the OLS regression results.[183] Also, applied post-estimation analyses show no threat of multicollinearity.[184]

To begin with model (1), unidimensional RPI seem to increase the difficulty level at which participants change from one bloc to another while the effect of multidimensional RPI appears to be negative as compared to the treatments that do not involve RPI. However, neither of the two forms of RPI significantly influence attention towards task dimensions (p=0.603, p=0.921). By contrast, the results show that the application of performance-based compensation is associated with a significant increase in the average difficulty level (p=0.023). Yet, the explanatory power of model (1) is relatively low and the model coefficients are overall not significantly different from zero (Prob>F=0.146, R^2=0.038).

In model (2), the addition of personal variables reveals that participants' risk aversion significantly lowers the difficulty level at which they change from one task bloc to another (p=0.001). On the other hand, it also indicates a significant positive effect of participants' preference for competition and of majoring in a math-related field of study (p=0.045, p=0.040). Apart from that, the effects of the two forms of RPI and of the compensation scheme do not considerably change as compared to model (1). In general, model (2) has a higher explanatory power than model (1) and is statistically significant (Prob>F=0.001, R^2=0.174).

Next, model (3) includes interaction terms between the dimensionality of RPI and the compensation scheme as well as the total rank of participants. Therewith, the predictive power of the model is increased (Prob>F=0.001, R^2=0.209). The results imply positive but insignificant coefficients for the effects of unidimensional and multidimensional RPI on the average difficulty level (p=0.113, p=0.329). Similar to model (1) and (2), the OLS estimates demonstrate a significant increase of the difficulty level under performance-based compensation (p=0.048). There seems to be no interaction between performance-based compensation and the provision of unidimensional RPI since the related coefficient is nearly zero (p=0.985). As opposed to that, the analysis reveals that multidimensional RPI overall decreases the average difficulty level under performance-based compensation in comparison to when participants receive no RPI. Although the coefficient is not significant, its size suggest that the interaction effect outweighs the positive influence indicated by the dummy variable for multidimensional RPI (p=0.246).

Model (3) additionally discloses that the interaction effects between the two forms of RPI and the overall ranking position of the participants are both negative. Yet, the negative interaction is only significant in case of unidimensional RPI (p=0.040, p=0.448). Notably, the coefficients show that the application of unidimensional RPI overall lowers the average difficulty level of low ranked participants as compared to the treatments which do not involve RPI. By contrast, the overall effect of multidimensional RPI on the average difficulty level remains to be positive with regard to low ranked participants. Finally, in model (3), the risk aversion of participants still significantly reduces the average difficulty level at which they change from one bloc to another (p=0.001). Besides, no other personal variable has a considerable influence on attention towards task dimensions.

[178] The related participant only reached a total number of points of 23 (IQR: - 1.80).

[179] Wald tests are applied in order to test for the equality of the coefficients (cf. Wald (1943)). Model (1): F=0.77, p=0.381; Model (2): F=1.77, p=0.186; Model (3): F=0.40, p=0.526.

[180] Wald test: F=0.72, p=0.398.

[181] Wald test: F=0.04, p=0.846.

[182] A comprehensive overview of the results of the median regressions for model (1) - (3) can be found in Online-Appendix 11.

[183] Breusch-Pagan test (cf. Breusch and Pagan (1979)) and White's test (cf. White (1980)) are conducted to check for the homoscedasticity of residuals. The Breusch-Pagan test (normality assumption removed) indicates heteroskedasticity for model (2) and model (3) ($\chi^2(1)$=6.09, p=0.014; $\chi^2(1)$=6.39, p=0.012).

[184] Estimated VIFs are constantly below the value 10 (cf. Marquaridt (1970)). VIF ranges are as follows: model (1): 1.00-1.33; model (2): 1.02-1.43; model (3): VIF=1.11-3.39.

Table 4: Dependent Variable Difficulty Level – Regression Results

Model (1) reports treatment group effects only. Model (2) additionally includes the control variables Gender (dummy variable, 1='male', 2='female'), Age (mean-centered), Math-related Field of Study (dummy variable, 1='Math-related Field of Study', 2='No Math-related Field of Study), Risk Aversion (range 0 - 10 where 0 indicates 'highly risk loving' and 10 indicates 'highly risk averse'), Preference for Competition (participants' responses on a seven-point Likert scale to the statement 'I perform better when I compete against someone rather than when I am the only one striving for a goal', 1='not at all', 7='to a great degree'). Model (3) adds interaction terms between the treatment groups for RPI and for compensation scheme as well as interaction terms between the treatment groups for RPI and a dummy variable for participants' position in the final ranking on overall performance (1='Rank 4 or 5', 2='Rank 1, 2 or 3'). Heteroskedasticity-consistent standard error estimators are implemented.
* Significant at 10%; ** Significant at 5%; *** Significant at 1% level.

Regressors (DV: Performance)	**Model (1) OLS Reg.**	**Model (2) OLS Reg.**	**Model (3) OLS Reg.**	**Model (4) OLS Reg.**
	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value	Coeff. (S.E.) p-value
Unidimensional RPI	0.1087 (0.2083) 0.6027	0.2514 (0.2052) 0.2224	0.4612 (0.2891) 0.113	
Multidimensional RPI	-0.02 (0.2009) 0.9214	-0.017 (0.2027) 0.9325	0.3228 (0.3298) 0.3294	
Perf.-based Compensation	0.394 (0.1713) 0.023**	0.4078 (0.1654) 0.015**	0.5566 (0.2795) 0.048**	
Gender		0.2393 (0.1689) 0.1588	0.2847 (0.1736) 0.1034	
Age		0.0349 (0.025) 0.1659	0.0337 (0.0246) 0.1727	
Math-related Field of Study		0.3977 (0.1915) 0.040**	0.3096 (0.2118) 0.1461	
Risk aversion		-0.176 (0.051) 0.001***	-0.174 (0.049) 0.001***	
Preference for Competition		0.0805 (0.0399) 0.045**	0.0648 (0.0409) 0.1151	
Unidimensional RPI x Perf.-based Compensation			0.0075 (0.397) 0.9849	

(Continued)

Table 4—continued

Regressors (DV: Performance)	Model (1) OLS Reg.	Model (2) OLS Reg.	Model (3) OLS Reg.	Model (4) OLS Reg.
Multidimensional RPI x Perf.-based Compensation			-0.47 (0.4033) 0.2461	
Unidimensional RPI x Low Rank			-0.553 (0.2669) 0.040**	
Multidimensional RPI x Low Rank			-0.251 (0.3295) 0.4477	
Constant	2.5726 (0.1499) 0.000***	2.7851 (0.4243) 0.000***	2.8106 (0.4456) 0.000***	
N	150	150	150	
Prob>F	0.1464	0.0006	0.0005	
R^2	0.0377	0.1737	0.2093	

Post-estimation analyses are applied in order to test whether the effects of unidimensional and multidimensional RPI significantly differ from each other in the presented models. Overall, no significant variation is found between the influences of the two forms of RPI on attention towards task dimensions under flat wage compensation.[185] Also, with regard to model (3), the interactions of unidimensional and multidimensional RPI with the compensation scheme[186] as well as with the overall ranking position of the participants[187] do not show significant differences.

5.1.3. Discussion

The effects of unidimensional RPI on performance and attention towards task dimensions (H1, H2)

The basis for the examination of the influence of the dimensionality of RPI on performance and attention towards task dimensions is laid by the first hypothesis H1 stating that unidimensional RPI increases performance in multidimensional task settings. The results of both the bivariate and the multivariate analyses fully support the formulated prediction. This substantiates the expected positive effects of RPI on an individual's motivation and learning which imply higher effort and the alignment of attention allocation towards overall task performance. Therefore, the findings of prior research on performance improvements under RPI appear to be generalizable to multidimensional task settings when RPI compares peer performance based on an overall performance measure that aggregates relevant task dimensions.

However, regarding the second hypothesis H2, the analysis does not confirm the expected performance increase when unidimensional RPI is provided together with performance-based instead of flat wage compensation. Particularly, neither an additive effect nor a positive interaction effect can be identified with respect to the motivational impact of performance-based compensation predicted by agency theory.[188] By contrast, results of the median regression suggest that unidimensional RPI significantly negatively interacts with the application of performance-based pay while its overall performance effect still stays considerably positively. With regard to the reasons of this discrepancy between the theoretical prediction and the experimental results, it is firstly of interest if there are unexpected differences in participants' attention towards task dimensions when unidimensional RPI is provided together with performance-based instead of flat wage compensation. Indeed, the bivariate analysis indicates a significant increase in the mean difficulty level under a performance-based contract. However, the corresponding interaction term in the regression analysis on attention towards task dimensions is nearly zero and insignificant. In this context, additional tests reveal that the result of the bivariate analysis is strongly driven by two observations in the data set

[185] Cf. Wald (1943); Wald test: Model (1): F=0.34, p=0.560; Model (2): F=1.70, p=0.194; Model (3): F=0.15, p=0.695.

[186] Wald test: F=1.36, p=0.245.

[187] Wald test: F=0.52, p=0.473.

[188] The graphical comparison of mean performance values in Figure 1 in section 5.1.1 originally suggested an additive effect. However, this effect on mean performance did not hold under an OLS regression which included control variables. The median performance values already indicated in 5.1.1 a decrease of performance which was confirmed in the median regression.

which relate to participants with an average difficulty level between 5.5 and 6.0.[189] Notably, the same participants also achieved above-average performance[190] and further analyses show that an increase in the average difficulty level under performance-based compensation can be mostly attributed to high performing participants.[191] Therefore, it appears that the introduction of performance-based compensation is not associated with an undesirable distortion in attention allocation under unidimensional RPI which might have impede performance improvements.

Importantly, the distribution of performance under unidimensional RPI suggests that the application of a performance-based contract has actually resulted in a positive effect with regard to high performing participants.[192] Also, the presence of extreme upward outlying observations under the performance-based scheme may be an indication that the predicted motivational effect of financial incentives has pushed participants with high ability to achieve maximum performance.[193] With respect to the participants with below-average performance, it should be emphasized that the nature of the applied task implies that the relationship between effort and performance is moderated by participants' cognitive abilities. Particularly, performance improvements through higher motivation under performance-based compensation are limited based on participants' ability to solve multiplication problems.[194] In this regard, performance-based compensation may have only positively affected performance of participants that did not receive RPI since the provision of RPI already improved performance under a flat wage contract which is why there might have been no much more room for improvement.

Additionally, it is worth mentioning that the application of a variable payment increases the performance pressure inherent in the setting which might be critical in conjunction with the competitive pressure induced by RPI. Notably, it could have resulted in a so called 'choking under pressure' effect which refers to the occurrence of inferior performance under pressure circumstances.[195] In this context, Beilock et al. (2004) demonstrate that pressure induced by monetary rewards harms performance in mathematical problem solving which requires heavy working memory such as multiplication tasks.[196] In the experimental setting, especially participants with below-average performance could have been affected by 'choking under pressure' since perceived pressure might increase through a low ranking position as well as the threat that monetary earnings are falling short of expectations.[197] However, additional analysis on the participants' mistake share and on the average time required to solve a problem of a certain difficulty level do not further underpin this supposition.[198]

Differences between the effects of unidimensional and multidimensional RPI on performance and attention towards task dimensions (RQ1a, RQ1b, RQ2a, RQ2b)

Besides the formulated hypotheses, the study has developed two-tailed research questions so as to address its objective of comparing the performance and attention allocation effects of unidimensional RPI with that of multidimensional RPI. Research question RQ2a is concerned with performance differences. Firstly, the bivariate as well as the multivariate analyses do not find a significant variation between the performance effects of the two forms of RPI. Yet, there is also no evidence for a significant positive performance effect of multidimensional RPI unless the regression analysis considers its interaction effects with the applied compensation scheme and the overall ranking position of the participants. Similar to unidimensional RPI, the results of the median regression reveal a significant negative interaction between multidimensional RPI and performance-based pay.[199] Importantly, however, the overall effect of multidimensional RPI under performance-based compensation is therefore nearly zero while unidimensional RPI still induces a remarkable performance increase. Regarding research question RQ2b, this complies with the finding of the bivariate analyses that median performance significantly differs between the two forms of RPI in the presence of a performance-based contract.

In order to interpret the described performance effects, potential differences in attention towards task dimensions between unidimensional and multidimensional RPI have to be considered. In relation to research question RQ1a, the analyses overall indicate that the average difficulty level under the two forms of RPI is not significantly different which

[189] When the observations are excluded from the data set, neither the Mann-Whitney test nor the Pearson Chi-Squared test indicate significant differences between the average difficulty level under perf.-based and flat wage compensation (p=0.22 / 0.39)

[190] The related participants achieved a total number of points of 602 (difficulty level 6.0) and 396 (difficulty level 5.7).

[191] When the sample is restricted to participants that perform above the mean of their treatment, mean /median difficulty levels are the following: Perf.-based: 3.82 / 3.56; Flat wage: 2.74 / 2.49. The difference is not significant (p=0.18 / 0.52).

[192] When the sample is restricted to participants that perform above the mean of their treatment, mean /median performance is the following: Perf.-based: 629.29 / 492; Flat wage: 453.42 / 389. The difference is not significant (p=0.12 /0.11).

[193] The extreme outlying observations relate to 2 participants with total points of 827 (IQR: 3.28) and 1312 (IQR: 6.79).

[194] Prior literature commonly argues that the positive performance effect of financial incentives predicted by agency theory is limited when tasks involve cognitive performance, cf. Bonner and Sprinkle (2002), p. 320; Camerer and Hogarth (1999); Rydval and Ortmann (2004).

[195] Cf. Baumeister (1984), p. 610; Ariely et al. (2009).

[196] Beilock et al. (2004) relates this finding to distraction theories which propose that pressure increases thoughts about the situation and its importance which reduce the capacity of the working memory that is available for the execution of the task.

[197] Commonly, participants in the experimental laboratory MELESSA expect to earn between 10 and 15 Euro per hour of participation. Since the show-up fee accounts for 4 Euro, participants had to achieve 300 points in order to assure that they earn at least 10 Euro in the experiment. (The earn- ings from the risk game are not disclosed until the end of the experiment and could have been in the minimum only 0.1 Euro.)

[198] Mistake share: (Number of incorrect answers)/ (Sum of number of correct and incorrect answers per difficulty level).

[199] As shown in section 5.1.2, the interaction effects are not significantly different from each other.

corresponds to the results on the performance effects. Further on, the introduction of performance-based compensation seem to differently affect the average difficulty level reached in the presence of unidimensional and multidimensional RPI. More concretely, the regression analysis reveals a negative interaction effect between multidimensional RPI and performance-based compensation and implies an overall decrease of the average difficulty level in comparison to the treatment that do not involve RPI. This contrasts with the very small and positive interaction between unidimensional RPI and performance-based pay. However, with respect to RQ1b, the effects of multidimensional RPI on attention towards task dimension are not statistically significant and neither is the difference between the interaction terms of the two forms of RPI with performance-based compensation.

From a theoretical perspective, it was outlined that multidimensional as compared to unidimensional RPI may differently affect attention allocation and performance because it do not limit peer comparisons to the overall performance measure which can promote learning but also motivate a distortion of attention. In the presence of performance-based compensation, the rewarding of overall performance were expected to weaken potential attention distortion effects and to strengthen the impact of learning benefits. Therefore, it surprises that negative performance effects of multidimensional RPI are present under performance-based compensation but not when a flat wage is applied. More in-depth analyses reveal that the decrease in performance between the two compensation contracts do not hold with regard to high performing participants which complies with the findings on unidimensional RPI.[200] Also, it can be shown that the increase in the average difficulty level suggested by the mean values, which opposes the indication of the median values and the regression results, can be attributed to those high performing participants.[201] In this context, a potential attention distortion towards the 'quantity' dimension of the task for participants performing below the treatment mean is further investigated. Notably, with respect to participants that achieved rank one or two in the 'quantity' ranking of their group, overall performance is considerably lower for multidimensional RPI under performance-based as compared to flat wage compensation. Also, the performance under both compensation contracts falls short to what is achieved by the top two performers of the 'quantity' ranking under unidimensional RPI.[202] This may indicate that participants provided with multidimensional RPI concentrate on the 'quantity' dimension of the task at the expense of overall performance, especially under performance-based compensation. However, the differences between the treatments are not significant and it appears that the performance decrease under multidimensional RPI in the presence of a performance-based contract might be also strengthened by further determinants.

As already mentioned before, the performance pressure inherent to a task setting is increased by the provision of RPI as well as by the application of a performance-based contract. Therefore, performance deteriorations of participants that are exposed to both forms of pressure might be linked to a 'choking under pressure' effect. In contrast to unidimensional RPI, additional analyses on the average time that participants under multidimensional RPI have required to solve a problem of a certain difficulty level provide some support for this supposition. Specifically, it can be shown that participants provided with performance-based compensation have spent more time on problems of all difficulty levels in comparison to those that received a flat wage. Notably, the differences between the compensation contracts are significant for difficulty level 1 and 2 when outliers are excluded from the sample.[203] Based on this finding, it may be argued that the provision of performance-based compensation increases pressure more under multidimensional RPI than under unidimensional RPI. From a theoretical perspective, this could be explained by the additional cognitive effort that is required when individuals have to solve trade-offs between the potentially conflicting motives induced by in the rewarding of overall performance and the facilitation of competition on dimension-specific performance measures.[204]

The effects of unidimensional and multidimensional RPI on performance and attention towards task dimensions of low ranked participants

Lastly, with regard to the overall performance effects of unidimensional and multidimensional RPI, the regression analysis reveals significant negative interactions between both forms of RPI and the overall ranking position of the participants.[205] In particular, the size of the effects of the OLS regression indi-

[200]When the sample is restricted to participants that perform above the mean of their treatment, mean / median performance is the following: Perf.-based: 517.63 / 447.5; Flat: 455.9 / 420.5. The difference is not significant (p=0.93 /1.00). Similarly to unidimensional RPI, the sample on perf.- based compensation also includes an extreme outlying observation. The related participant achieved a total number of points of 1217 (IQR: 6.38).

[201]When the sample is restricted to participants that perform above the mean of their treatment, mean /median difficulty levels are the following: Perf.-based: 3.23 / 2.76; Flat: 3.02 / 2.64. The dif- ference is not significant (p=0.96 / 1.00). The distribution under perf.-based compensation includes two outlying observation related to participants with an average difficulty level of 4.71 (IQR: 2.16) and 6.0 (IQR: 3.81). The same participants also represent the two outlying observa- tions on performance. However, performance as well as the average difficulty are for both partici- pants that perform above and below the mean still considerably higher under unidimensional RPI and perf.-based compensation.

[202]Mean and median performance is as follows: Multidimensional RPI: 1) Flat: 322.7 / 327.5, Perf.- based: 296.9 / 303.5; Unidimensional RPI: 1) Flat: 363 / 354, Perf.-based: 354 / 311. The differences are not significant (all p>0.37).

[203]Mean and median time values: Difficulty level 1) Flat: 9.42 / 8.23; Perf.-based: 11.36 / 10.48 (p=0.24 / 0.04); Difficulty level 2) Flat: 10.59 / 10.25; Perf.-based: 13.64 / 12.57 (p=0.04 / 0.14). The differences in the time values are not expected to be driven by different abilities since the treatments do not differ in their share of students majoring in a math-related field of study and since the first two difficulty levels only require basic math skills. There are no significant differ- ences in the average mistake shares per difficulty level.

[204]Cf. Payne et al. (1993), p. 30; Hogarth (1987).

[205]As shown in section 5.1.2, the interaction effects are not significantly different from each other. It should be emphasized that low ranked participants are not necessarily below-average performers of their treatment since the participants compete in groups of five while each treatment is applied to

cate that the performance of low ranked participants overall decreases when they are provided with either form of RPI. In the median regression, this only applies to multidimensional RPI.[206] In interpreting these findings, it should be referred to the insight of social comparison theory that comparisons with superior others can be on the one the hand inspiring and spur the motivation for performance improvement but may also adversely affect self-evaluation and promote negative feelings of inferiority. As outlined earlier in section 2.1.1, the latter is expected to occur when a comparison implies a contrastive evaluation of the comparison target which emphasizes the individual's separation toward it.[207] More concretely, comparison standards are perceived to be not attainable which might have applied to the participants of the experiment which experienced a ranking position below the top three ranks.

In the described context, peer comparisons can firstly decrease task motivation and provoke defensive reactions which can result in lower effort and deteriorating performance.[208] Secondly, threatening upward comparisons can lower self-efficacy beliefs and therefore cause a devotion of cognitive effort towards activities that are unrelated to task execution such as worrying about the own performance and interpreting the ranking outcome.[209] Here, prior experimental studies show that the behavioural responses can also include inconsistent changes of task strategies and the adoption of risky strategies.[210] With regard to the performance decrease of the low ranked participants in the conducted experiment, some evidence is found with regard to the second explanation. Particularly, the responses to the post-experimental questionnaire indicate that low ranked as compared to high ranked participants were more nervous about their rank and perceived that thinking about their rank interfered to a greater extent with their performance (p=0.02, p=0.08).[211] Also, it appears that low ranked participants changed their task strategy more often during the experiment since the number of changes in the average difficulty level between the rounds is significantly greater in comparison with high ranked participants (p=0.01).[212] Moreover, it can be shown that the share of incorrect answers on the total number of answers is significantly higher for low ranked participants (p=0.00).[213] In this context, they have also required significantly more time to solve problems of difficulty level one to five (all p<0.00).[214] However, this can be as well an indication for decreases in the effort level due to a loss of motivation.

As mentioned before, the regression results indicate that the negative performance effect of the provision of RPI on low ranked participants appears to be specifically emphasized under multidimensional RPI. This result may be interpreted in two different ways. First, participants that are provided with multidimensional RPI have the opportunity to reduce the self-threating effect induced by a negative social comparison on overall performance through the affirmation of their competencies in one of the task dimensions. Therefore, the predictions derived from self-affirmation theory with regard to potential attention distortion effects under multidimensional RPI might be especially relevant for low ranked participants.[215] In this regard, the interaction term included in the regression analysis on attention towards task dimensions indicates a decrease in the average difficulty level of low ranked participants. However, the effect is not significant and additional analyses show that there are is also no difference between the two forms of RPI regarding the number of low ranked participants that achieved a high rank in the 'quantity' or the 'value' dimension of the task.[216] Notably, the regression results disclose a significant negative interaction between unidimensional RPI and the average difficulty level of low ranked participants. This might be explained by decreased motivation and a lower effort level since tasks of lower difficulty require less cognitive effort and the provision of unidimensional RPI does not involve potential incentives for a distortion of attention towards one task dimension. Indeed, it can be demonstrated that the significant effect is driven by two participants with a considerably low respectively downward outlying overall performance.[217]

25 participants in total. Therefore, the study compares the overall effect of the two forms of RPI with regard to low ranked participants and disentangles effects between above and below-average performers only in the context of the application of different compensation schemes.

[206]As mentioned before, the data on unidimensional RPI include a downward outlying observation which relates to a low ranked participant (performance: 23 points, IQR: -1.80).

[207]Cf. Festinger (1954); Suls et al. (2002); Corcoran et al. (2011); Tesser (1988).

[208]Cf. Corcoran et al. (2011); Stapel and Schwinghammer (2004). Defensive reactions may include e.g. that individuals question the relevance of the performance dimension, search for differences be- tween themselves and the comparison target or try to self-affirm their competencies in another area.

[209]Cf. Bandura (1977b); Bandura/Wood (1989).

[210]Cf. Hannan et al. (2008); Camerer and Hogarth (1999).

[211]Participants had to indicate rank nervousness and interference using a seven-point Likert scale (1=not at all, 7=to a great degree). The questions are adapted from Tafkov (2013). See Online-Appendix 5 for the questions of the post-experimental questionnaire. Based on the answers, continuous varia- bles are created. Mean values: Rank nervousness - Low rank: 3.18 - High rank: 2.35; Rank inter- ference - Low rank: 2.83 - High rank: 2.15. P-values are derived from a Mann-Whitney test. The comparison is limited to participants that received RPI (unidimensional or multidimensional RPI). There are no significant differences in the responses of low and high ranked participants that were not provided with RPI (and therefore did not know their ranking position).

[212]Mean number of changes in the average difficulty level: Low rank: 1.13 – High rank: 0.63. P-values are derived from a Mann-Whitney test. The changes are counted from one period to another and relate to an increase or a decrease of the average difficulty level of at least one level (e.g. a change from difficulty level 2.0 to 3.1 between period 2 and 3 is counted as one change).

[213]Mean values on mistake share: Low rank - 18.75% - High rank: 27.56%. P-values are derived from a Mann-Whitney test.

[214]Low ranked participants have also required more time to solve problems of difficulty level six but the difference is not significant since there are only 6 observations as compared to 22 observations for high ranked participants. P-values are derived from a Mann-Whitney test.

[215]Cf. Steele (1988); Aronson et al. (1999).

[216]The tests were conducted with different definitions of 'high rank': Rank 1; Rank 1 or 2; Rank 1, 2 or 3.

[217]The mean difficulty level of low ranked participants under unidimen-

Apart from that, multidimensional RPI may have an increased negative impact on the performance of low ranked participants as it can strengthen adverse effects on self-evaluation. More concretely, it potentially involves negative social comparison experiences in different competence areas. This can be illustrated by the finding that nearly all of the overall low ranked participants were also confronted with a low rank regarding their dimension-specific performance.[218] Therefore, it could be argued that multidimensional RPI as compared to unidimensional RPI potentially reinforces negative effects on task motivation and self-efficacy beliefs. Particularly, low ranked participants may have exerted less effort because of decreased motivation since there are no significant differences between the two forms of RPI with regard to strategy changes, mistake share, rank nervousness and interference as well as time required to solve problems of certain difficulty levels.

5.2. Supplemental analyses

5.2.1. Development of performance and attention towards task dimensions over time

In the previous chapter, tests of hypotheses and research questions have used data pooled across the six rounds of the experiment. In order to extend the provided insights on the effects of the experimental manipulations, this section considers how the primary dependent variables performance and attention towards task dimensions have developed over time.

Figure 3 shows the time path of the mean performance of participants by round in all treatments. The graph illustrates that performance differences based on the presence and the dimensionality of RPI already exist in the first round before the participants in the related treatments actually receive RPI. This finding is consistent with previous experimental studies which show that individuals anticipate the potential effect of RPI on their self-evaluation and therefore respond accordingly.[219] In particular, when the performance means are pooled across compensation contracts, they correspond to the results on overall mean performance. Participants' number of points is significantly higher when they are provided with unidimensional RPI (p=0.08) while the positive effect of multidimensional RPI is not significant (p=0.25).[220] However, initially there seem to be no differences between the applied compensation schemes except for a slightly higher level of performance when multidimensional RPI is combined with a flat wage instead of performance-based contract. The moderating influence of the compensation contract under multidimensional RPI appears to strengthen in the following rounds which implies that mean performance is constantly lower in the presence of a performance-based as compared to a flat wage scheme.[221] This insight underpins the results of the previous analyses regarding a negative interaction effect between multidimensional RPI and performance-based compensation.

In contrast to the described differences between the compensation contracts under multidimensional RPI, Figure 3 suggests that performance is after the first round consistently higher in the presence of a performance-based scheme when participants receive unidimensional RPI or no RPI. Further analyses indicate that median performance under unidimensional RPI is higher for flat wage compensation in the first two rounds of the experiment which may explain the negative interaction effect implied by overall median performance and the regression results.[222] The evolution of performance therefore indicates that negative influences of performance-based compensation are relevant in the early rounds while its predicted motivational effects appear to manifest itself in higher performance towards the end of the experiment. This finding could be explained by an increase in practice on the basic multiplication tables that are needed to solve problems of all difficulty levels since this potentially reduces performance declines because of 'choking under pressure'. Specifically, Beilock et al. (2004) argues that 'choking under pressure' may be eliminated in mathematical problem solving as individuals gain practice and can therefore retrieve partial solutions from memory.[223] Importantly, yet, such a countervailing effect of practice on potential 'choking under pressure' in the presence of performance-based compensation cannot be observed with regard to multidimensional RPI.

Overall, the development of performance under unidimensional RPI and performance-based compensation is characterized by some volatility over the first three rounds and a steady increase with regard to the following rounds. Under the flat wage scheme, the development appears to more stable with a slight rise of performance over time. By contrast, the graphs show considerable volatility in performance over all six rounds when participants do not receive RPI or are provided with multidimensional RPI. Remarkably, for both manipulations, the evolution under the two compensation contracts follows the same pattern even though on different total performance levels. Furthermore, there is no considerable directional difference in the evolution of performance under the treatments which do not involve and those that apply multidimensional RPI. The only exception applies to the last round in which the performance under multidimensional

sional RPI equals 2.53. In contrast, the two participants reached an average difficulty level of 1.86 (performance: 87 points) and 1.5 (performance: 23 points, IQR: -1.80).

[218] 16 out of 20 overall low ranked participants under multidimensional RPI were also low ranked on either the 'quantity' or the 'value' dimension of the task (rank 4 or 5). Two low ranked participants even received a low rank on both task dimensions.

[219] Cf. Tafkov (2013); Hannan et al. (2013).

[220] P-values are derived from a Mann-Whitney test. The same result can be obtained from a Pearson Chi-Squared test with regard to median performance (p=0.07, p=0.23). No significant difference in performance can be found between unidimensional and multidimensional RPI in the first round.

[221] Based on the Mann-Whitney test, there is no significant difference in performance under the two compensation contracts in all rounds. Yet, with regard to median performance, the Pearson Chi-Squared test indicates a significant positive effect under a flat wage in round three and four (p=0.05, p=0.09).

[222] There is no significant difference in performance between the two compensation contracts in all rounds.

[223] Cf. Beilock et al. (2004).

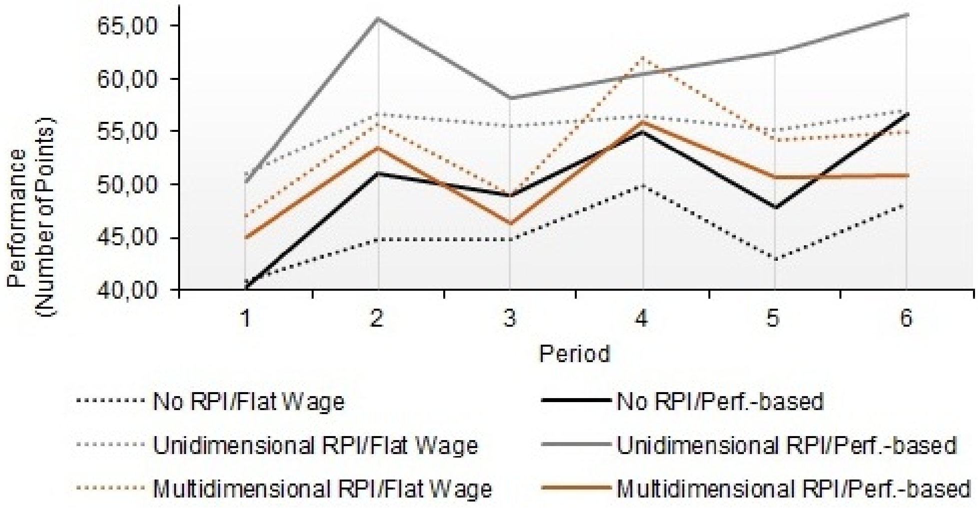

Figure 3: Development of Performance over Periods (Means by Treatment)

slightly decreases while there is a substantially rise for the performance of participants that are not provided with RPI. This result also holds for median performance and can indicate a loss of motivation for participants under multidimensional RPI. Since the provided rankings did not compare current period performance but cumulative performance, they may have not seen the chance to improve their position in the ranking in the last period. Lastly, the development of performance in all treatments reflects an overall increase from the first to the last period which can be related to training and learning effects.[224]

Figure 4 displays the evolution of the mean values per treatment for participants' average difficulty level over all six rounds of the experiment. Broadly, the graphs indicate for nearly all treatments that the average difficulty level constantly fluctuate around a certain baseline with overall changes of no more than 0.58 difficulty level. The only exception is the treatment that does not involve RPI and applies a flat wage since it shows a decrease of 0.84 difficulty levels from the first to the fourth round. While the average difficulty level is initially similar between the two compensation contracts when participants do not receive RPI, this reduction causes an increasing difference since the average difficulty level under performance-based compensation shows a relatively low fluctuation.[225] The development partly complies with the results on performance which show similarity in the first round and a consistently higher level of performance under performance-based compensation in the following rounds. Specifically, it appears that participants provided with a flat wage cannot realize similar performance improvements with the lower difficulty level as those that receive performance-based contract and stay with the higher difficulty level. A possible reason for the decrease of the average difficulty level under flat wage compensation may be the lower level of motivation since solving problems of lower difficulty level requires less cognitive effort.

With regard to the evolution of attention towards task dimensions under unidimensional RPI, Figure 4 first of all shows that the average difficulty level of participants under performance-based compensation is over the whole time the highest among all treatments. Thus, it is from the beginning constantly higher than the average difficulty level under unidimensional RPI and flat wage compensation even though first round performance is similar between both compensation contracts.[226] Notably, the average difficulty level under both compensation schemes fluctuates in the same pattern and seem to be unrelated to the presented stable performance development in the latter rounds of the experiment. Unlike the development under unidimensional RPI, the average difficulty level of participants provided with multidimensional RPI evolves around a comparable baseline under the two compensation contracts. However, while the mean of changes in the average difficulty level over time is lowest for participants provided with a flat wage as compared to all other treatments, there is a considerable volatility under a performance-based scheme.[227] The differences in the development of the average difficulty level are not reflected in the previously shown similar pattern of the evolution of performance under the two compensation contracts. Nevertheless, they might be related to the overall higher performance level under a flat wage scheme. Finally, a comparison between the evolution of the average difficulty level over time

[224] The greatest increases occurred under No RPI/Perf.-based and Unidimensional RPI/Perf.-based. The differences in the mean performance values between the first and last period equal 16.36 and 15.96. For the other treatments, the differences are all below 7.96.

[225] The average difficulty level is significantly higher in round four (p=0.05), five (p=0.05) and six (0.01). The Pearson Chi-Squared test indicates a median difference in round six (p=0.05).

[226] The average difficulty level is significantly higher in round one (p=0.04) and three (p=0.04). The Pearson Chi-Squared test indicates a median difference in round three (p=0.05).

[227] Mean number of changes in the average difficulty level: 1.0. In the remaining treatments, the mean is below 0.88.

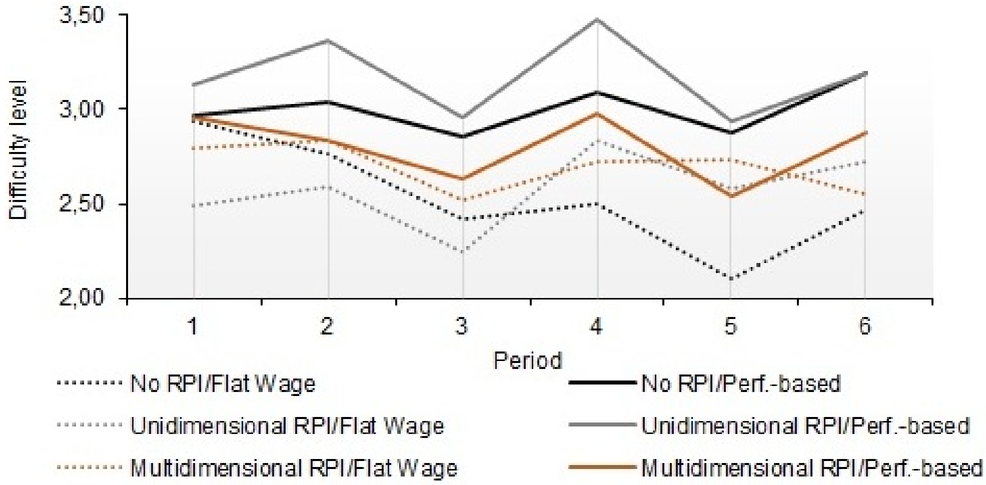

Figure 4: Development of Difficulty Level over Periods (Means by Treatment)

under multidimensional RPI as compared to unidimensional RPI provides no substantiation on differences in learning effects. In this regard, especially the behavioural responses to the provision of RPI in the first round may be relevant which though do not indicate improved adjustments of the average difficulty level under multidimensional RPI.

5.2.2. Post-Experimental questionnaire data

In addition to the observed primary dependent variables of the experiment, post-experimental questionnaire data are used to provide some further insights into the effects of unidimensional and multidimensional RPI.[228] Specifically, it is examined how participants' ranking position relates to their indicated feelings of pride and shame and whether there is a connection between the dimensionality of RPI and statements regarding the development and the use of task-specific strategies.

In the course of the discussion of the results in section 5.1.3, it was argued that social comparison information can adversely affect self-evaluation of low ranked participants and foster negative feelings of inferiority based on a contrastive evaluation of the comparison target. In this context, the data of the post-experimental questionnaire allow to investigate whether participants' position in the rankings created feelings of pride and shame. Using a seven-point Likert scale, participants had to assess how pride they felt about their own performance (1=not at all, 7=to a great degree).[229] With regard to the theoretical predictions, it is expected that the continuous variable for participants' total rank is negatively related to the variable that reflects pride since a smaller rank number constitutes better relative performance. Additional analyses support this supposition since they show that total rank and pride are significantly negatively correlated for both participants that were provided with unidimensional and multidimensional RPI (r= -0.69, p=0.00; r= -0.75, p=0.00).[230] It appears that there is a stronger negative correlation for multidimensional RPI which could be an indication for a strengthening of adverse effects on self-evaluation when participants potentially experience negative social comparison experiences in different competence areas. However, a test on the equality of the correlation coefficients reveals that there is no significant difference. Importantly, no correlation is found for participants that did not receive any form of RPI and were therefore not aware about differences between their own performances and that of their peers.

Apart from that, the post-experimental questionnaire was also used to ask participants about the task strategy that they adopted during the experiment as well as the usefulness of the provided ranking information for finding an optimal strategy. In this regard, it has been hypothesized in the analysis on the behavioural effects of RPI in section 2.1 that multidimensional RPI is more beneficial for learning processes on task strategies than unidimensional RPI. However, so far, the analyses on the observed primary dependent variables did not indicate differences in learning effects between the two forms of RPI. Notably, the post-experimental questionnaire data suggest that multidimensional RPI in general increased the application of a task-specific strategy as compared to unidimensional RPI. More concretely, the number of participants that constituted in the questionnaire that they have not followed any strategy during task execution is significantly higher under unidimensional RPI (p=0.09).[231] Nevertheless, the questionnaire data show that participants which received

[228]See Online-Appendix 5 for the questions of the post-experimental questionnaire.

[229]The question is adapted from Tafkov (2013).

[230]Regarding the additional rankings provided under multidimensional RPI, a significant correlation can be found between pride and the final position in the 'value' ranking (r=-0.47, p=0.00). How- ever, it appears that this is driven by a significant positive correlation between overall rank and 'value' rank (r=0.48, p=0.00).

[231]Participants where asked which strategy they have adopted during task execution and had to choose between the options 'Value Strategy', 'Quantity

multidimensional RPI did not perceive the ranking information to be more useful for strategy development than those that were provided with unidimensional RPI.[232] In addition, there appear to be no differences in the evaluation between the two compensation schemes that were applied with multidimensional RPI. Overall, the participants' average assessment of the usefulness equaled 3.45 (s.d.=2.01) on a seven point Likert scale (1=not at all; 7=to a great degree).

6. Conclusion

The provision of RPI is an important ingredient in the 'gamification' of tasks and is commonly used by firms to increase the performance of their employees. In case employees have to fulfil tasks that involve multiple dimensions, firms have to decide on the dimensionality of RPI and can basically choose between unidimensional and multidimensional RPI. However, little is known about the potential cost and benefits of the provision of different forms of RPI in multidimensional task settings in which overall performance is not only determined by the motivated level of effort but also by the induced attention towards task dimensions. Against this background, the present study has conceptually analysed behavioural effects of unidimensional and multidimensional RPI under different compensation schemes. On this basis, it has applied a controlled laboratory experiment to empirically test the influence of the dimensionality of RPI on performance and attention towards task dimensions.

In sum, the study demonstrates that solely the provision of unidimensional RPI improves performance in multidimensional task settings while the effects for multidimensional RPI are small and insignificant. Importantly, both unidimensional and multidimensional RPI imply a negative effect on the performance of participants with a low position in the provided ranking on overall performance. Additionally, the application of performance-based compensation negatively moderates the performance impact of both forms of RPI which though seem to be especially critical for multidimensional RPI. In this regard, multidimensional RPI induces a significant performance decrease compared to unidimensional RPI in the presence of a performance-based contract. The findings provide some indication that a distortion of attention toward the 'quantity' dimension of the applied experimental task may have caused the negative performance effect. Notably, the experimental results do not indicate increased learning effects regarding a task-specific strategy under multidimensional as compared to unidimensional RPI. Overall, no significant differences in attention towards task dimensions between the two forms of RPI can be proved.

The findings of the study extend the stream of literature that investigates the motivational influence of RPI in task settings with one relevant performance dimension since it enhances the understanding of the effects of RPI in multidimensional task setting. Furthermore, it involves practical implications for the use of RPI in firms which employ multidimensional tasks and have no certain preferences on employees' attention towards task dimensions but rather aim to maximize overall performance. Based on the study's results, it appears to be not beneficial to rank employees based on their dimension-specific performance. In this regard, features of 'gamification' platforms like the possibility to provide employees with symbolic badges for the mastery of different performance dimensions should be used with care.[233] Also, the study emphasizes that firms should not separate the decisions on the implementation of RPI and the use of financial incentives but have to jointly consider possible effects on the motivation of their workforce. Moreover, it is of special relevance regarding the growing adaption of 'gamification' platforms that firms are aware that the provision of RPI can come at the cost of demotivating low ranked employees.

The results of the present study are subject to a number of limitations. As already outlined in chapter 4.1, the generalizability of findings obtained from laboratory experiments is an essential concern.[234] First, results are drawn from a small and homogenous sample of 150 student participants. Second, anonymity was preserved during the experiment and therefore participants did not know how the fictional names used in the rankings matched to the actual persons in the laboratory. This may have lowered the strength of social comparison involvement compared to a workplace context in which colleagues know each other and where 'gamification' leaderboards commonly show photos of employees next to their rank.[235] Also, a lack of identification and psychological closeness to the comparison targets may have promoted the salience of a personal self in opposition to a social self which can enforce negative feelings and demotivation for low ranked participants.[236] Third, the applied experimental task enables a clear test of the underlying theory but is not reflective for multidimensional tasks in the real world as it does not capture all higher-order capabilities and task strategy considerations that might be required from employees to perform well in their jobs. Lastly, in general, it has to be emphasized that the experimental data include various (extreme) outlying observations which reduce the robustness of the results.

Overall, the findings of the study suggest various avenues for further research. Future investigations could shed light on the indicated moderating influence of the applied compensation scheme on the performance effects of RPI in multidimensional task settings. In this context, it seems to be of interest to explore further determinants of varying performance under the two forms of RPI apart from potential

Strategy', 'No Strategy' or 'Other Strategy' In case of a choice of 'Other Strategy', the applied strategy had to be specified. Percentage of participants that chose 'No Strategy': 1) Unidimensional RPI: 20%; 2) Multidimensional RPI: 8%.

[232] There is also no correlation between the perceived usefulness of the ranking information and the performance of the participants under the two forms of RPI.

[233] Cf. Werbach and Hunter (2012); The Economist (2012b)

[234] Cf. Loewenstein (1999); Schnell et al. (2008).

[235] Cf. Werbach and Hunter (2012); Bunchball.

[236] Cf. Stapel and Koomen (2005); Corcoran et al. (2011).

differences in attention towards task dimensions. Especially, potential 'choking under pressure' in the presence of multidimensional RPI may receive attention in connection with the cognitive effects of trade-offs faced by employees when financial rewards are tied to overall performance while competition is facilitated on dimension-specific performance measures. As it was outlined earlier, 'choking under pressure' is a frequent phenomenon in mathematical problems solving and therefore, it appears to be important to additionally test effects with other types of tasks. Furthermore, future research could investigate in greater depth whether multidimensional RPI may enhance adverse performance effects for low ranked employees as it can cause negative social comparison experiences in different competence areas. Finally, since the participants of the conducted experiment were only informed about performance ranks but not about performance levels of their group members, it would be interesting how such additional information may influence performance as well as attention allocation under unidimensional and multidimensional RPI.

References

Acock, C. A. *A Gentle Introduction to Stata*. College Station, 2006.

Antle, R. and Smith, A. An Empirical Investigation of the Relative Performance Evaluation of Corporate Executives. *Journal of Accounting Research*, 24(1):1–39, 1986.

Ariely, D., Gneezy, U., Loewenstein, G., and Mazar, N. Large stakes and big mistakes. *The Review of Economic Studies*, 76(2):451–469, 2009.

Aronson, J., Cohen, G., and Nail, R. P. Self-affirmation theory: An update and appraisal. In *Progress on a pivotal theory in social psychology*, pages 127–147. Eddie Harmon-Jonesand Judson Mills, Washington D. C., 1999.

Azmat, G. and Iriberri, N. The importance of relative performance feedback information: Evidence from a natural experiment using high school students. *Journal of Public Economics*, 94:435–452, 2010.

Baiman, S. Agency Research in Managerial Accounting: A Survey. *Journal of Accounting Literature*, 1:154–213, 1982.

Baiman, S. Agency research in managerial accounting: A second look. *Organizations and Society*, 15(4):341–371, 1990.

Baker, G. Distortion and risk in optimal incentive contracts. *Journal of human resources*, pages 728–751, 2002.

Bandura, A. *Social learning theory*. New York, 1977a.

Bandura, A. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2):191–215, 1977b.

Bandura, A. and Jourden, F. J. Self-regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology*, 60(6):941–951, 1991.

Bandura, A. and Wood, R. Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology*, 56(5):805–814, 1989.

Baumeister, R. F. Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of personality and social psychology*, 46(3):610–620, 1984.

Beach, S. R. H. and Tesser, A. Self-esteem and the extended self-evaluation maintenance model. In Michael, M., editor, *Efficacy, agency, and self-esteem*, pages 145–170. Kernis, New York, 1995.

Beilock, S. L., Kulp, C. A., Holt, L. E., and Carr, T. H. More on the Fragility of Performance: Choking Under Pressure in Mathematical Problem Solving. *Journal of Experimental Psychology*, 133(4):584–600, 2004.

Berger, L., Klassen, K., Libby, T., and Webb, A. Evidence on the Relation between Incentives and Performance in Repeated Tournaments. *Working paper Brock University and University of Waterloo*, 2011.

Blanes i Vidal, J. and Nossol, M. Tournaments without prizes: Evidence from personnel records. *Management Science*, 57(10):1721–1736, 2011.

Bonner, S. E. and Sprinkle, G. B. The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research. *Accounting, Organizations and Society*, 27(4):303–345, 2002.

Boudreau, K., Helfat, C., Lakhani, K. R., and Menietti, M. E. *Field evidence on individual behavior and performance in rank-order tournaments*. Harvard Business School Working PaperBoudreau2012, 2012.

Bourdieu, P. *La distinction: Critique sociale du jugement*. Paris, 1979.

Bracha, A. and Fershtman, C. Competitive Incentives: Working Harder or Working Smarter? *Management Science*, 59(4):771–781, 2013.

Breusch, T. S. and Pagan, A. R. A simple test for heteroscedasticity and random coefficient variation. *Journal of the Econometric Society*, pages 1287–1294, 1979.

Brickman, P. and Bulman, R. J. Pleasure and pain in social comparison. In *Social comparison processes: Theoretical and empirical perspectives*, volume 149, page 186. 1977.

Brown, D. J., Ferris, D. L., Heller, D., and Keeping, L. M. Antecedents and consequences of the frequency of upward and downward social comparisons at work. *Organizational Behavior and Human Decision Processes*, 102(1):59–75, 2007.

Brown, G. W. and Mood, A. M. On median tests for linear hypotheses. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. 1951.

Brown, J. L., Fisher, J. G., Sooy, M., and Sprinkle, G. B. The effect of rankings on honesty in budget reporting. *Accounting, Organizations and Society*, 39(4):237–246, 2014.

Brüggen, A. and Moers, F. The role of financial incentives and social incentives in multi-task settings. *Journal of Management Accounting Research*, 19(1):25–50, 2007.

Bunchball. Gamification Case Studies & Customers. URL http://www.bunchball.com/customers. Downloaded: 27.04.2014.

Buunk, B. P. and Mussweiler, T. New directions in social comparison research. *European Journal of Social Psychology*, 31(5):467–475, 2001.

Camerer, C. F. and Hogarth, R. M. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty*, 19(1-3):7–42, 1999.

Carmona, C., Buunk, A. P., Dijkstra, A., and Peiro, J. M. The relationship between goal orientation, social comparison responses, self-efficacy, and performance. *European Psychologist*, 13(3):188–196, 2008.

Castellan, J. N. and Swaine, M. Long-term feedback and differential feedback effects in nonmetric multiple-cue probability learning. *Behavioral Science*, 22(2):116–128, 1977.

Cervone, D., Peake, P. K., and Philip, K. Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology*, 50(3):492, 1986.

Charness, G. and Grosskopf, B. Relative payoffs and happiness: an experimental study. *Journal of Economic Behavior & Organization*, 45(3):301–328, 2001.

Charness, G. and Rabin, M. Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*, 117(3):817–869, 2002.

Charness, G., Masclet, D., and Villeval, M. C. The dark side of competition for status. *Management Science*, 60(1):38–55, 2013.

Chung, K. H. and Vickery, D. W. Relative effectiveness and joint effects of three selected reinforcements in a repetitive task situation. *Organizational Behavior and Human Performance*, 16(1):114–142, 1976.

Collins, R. L. For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 1:51–69, 1996.

Collins, R. L. Among the Better Ones. In *Handbook of Social Comparison*, pages 159–171. Jerry Sulsand Ladd Wheeler, 2000.

Colombino, T., Castellani, S., Grasso, A., and Willamowski, J. Agentville: Supporting Situational Awareness and Motivation in Call Centres. In *From Research to Practice in the Design of Cooperative Systems: Results and Open Challenges*, pages 291–307. Julie Dugdale, 2012.

Congleton, R. D. Efficient status seeking: externalities, and the evolution of status games. *Journal of Economic Behavior & Organization*, 11(2):175–790, 1989.

Corcoran, K., Crusius, J., and Mussweiler, T. Social comparison: Motives, standards, and mechanisms. *Theories in social psychology*, pages 119–139, 2011.

Croson, R. and Gneezy, U. Gender differences in preferences. *Journal of Economic literature*, pages 448–474, 2009.

Datar, S., Kulp, S. C., and Lambert, R. A. Balancing performance measures. *Journal of Accounting Research*, 39(1):75–92, 2001.

Deci, E. L. and Ryan, R. M. *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, 1985.

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference*, pages 9–15. Envisioning Future Media Environments, 2011.

Deutsch, M. A theory of cooperation and competition. *Human Relations*, 2: 129–152, 1949.

Eisenhardt, K. M. Agency theory: An assessment and review. *Academy of management review*, 14(1):57–74, 1989.

Eriksson, T., Poulsen, A., and Villeval, M. C. Feedback and incentives: Experimental evidence. *Labour Economics*, 16(6):679–688, 2009.

Falk, A. and Fehr, E. Why labour market experiments? *Labour Economics*, 10(4):399–406, 2003.

Falk, A. and Heckman, J. J. Lab experiments are a major source of knowledge in the social sciences. *Science*, 23:535–538, 2009.

Falk, A. and Ichino, A. Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24(1):39–57, 2006.

Falk, A., Dohmen, T., and Sunde, U. Kontrolliert und repräsentativ: Beispiele zur Komplementarität von Labor- und Felddaten. *Perspektiven der Wirtschaftspolitik*, 10:54–74, 2009.

Fehr, E. and Schmidt, K. M. Fairness and Incentives in a Multi-task Principal-Agent Model. *The Scandinavian Journal of Economics*, 106(3):453–474, 2004.

Feltham, G. A. and Xie, J. Performance measure congruity and diversity in multi-task principal/agent relations. *Accounting Review*, pages 429–453, 1994.

Feltovich, N. What's to Know about Laboratory Experimentation in Economics? *Journal of Economic Surveys*, 25(2):371–379, 2011.

Festinger, L. Informal social communication. *Psychological Review*, 57(5): 271–282, 1950.

Festinger, L. A theory of social comparison processes. *Human relations*, 7 (2):117–140, 1954.

Field, A. and Hole, G. J. *How to design and report experiments*. London, 2002.

Fischbacher, U. z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.

Frey, B. S. Awards as compensation. *European Management Review*, 4(1): 6–14, 2007.

Garcia, S. M. and Tor, A. Rankings, standards, and competition: Task vs. scale comparisons. *Organizational Behavior and Human Decision Processes*, 102(1):95–108, 2007.

Garcia, S. M. and Tor, A. The n-effect more competitors, less competition. *Psychological Science*, 20(7):871–877, 2009.

Garcia, S. M., Tor, A., and Gonzalez, R. Ranks and rivals: A theory of competition. *Personality and Social Psychology Bulletin*, 32(7):970–982, 2006.

Ginsburgh, V. Awards, success and aesthetic quality in the arts. *Journal of Economic Perspectives*, 17(2):99–111, 2003.

Gist, M. E. and Mitchell, T. R. Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management review*, 17(2): 183–211, 1992.

Goethals, G. R. and Darley, J. M. Social comparison theory: An attributional approach. In *Theoretical and empirical perspectives*, pages 259–278. Social comparison processes, 1977.

Greiner, B. An Online Recruitment System for Economic Experiments. In Macho, K. K. V., editor, *Forschung und wissenschaftliches Rechnen*, pages 79–93. Ges. fuer Wiss. Datenverarbeitung, Goettingen, 2004.

Griffin-Pierson, S. The Competitiveness Questionnaire: A Measure of Two Components of Competitiveness. *Measurement and Evaluation in Counseling and Development*, 23(3):108–115, 1990.

Guimond, S. *Social Comparison and Social Psychology: Understanding Cognition, Intergroup Relations, and Culture*. Cambridge, 2006.

Hannan, L. R., Mcphee, G. P., Newman, A. H., and Tafkov, I. D. *The Effect of Relative Performance Information Detail and Time Horizon on Effort Allocation in a Multi-Task Environment*. Working paper, 2014.

Hannan, R. L., Krishnan, R., and Newman, A. H. The Effects of Disseminating Relative Performance Feedback in Tournament and Individual Performance Compensation Plans. *Accounting Review*, 83(4):893–913, 2008.

Hannan, R. L., Mcphee, G. P., Newman, A. H., and Tafkov, I. D. The Effect of Relative Performance Information on Performance and Effort Allocation in a Multi-Task Environment. *Accounting Review*, 88(2):553–575, 2013.

Hattie, J. and Timperley, H. The power of feedback. *Review of educational research*, 77(1):81–112, 2007.

Hibbard, D. R. and Buhrmester, D. Competitiveness, gender, and adjustment among adolescents. *Sex Roles*, 63(5-6):412–424, 2010.

Hoaglin, D. C., Iglewicz, B., and Tukey, J. W. Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association*, 81(396):991–999, 1986.

Hogarth, R. M. *Judgement and choice*. The psychology of decision, 1987.

Holmstrom, B. Moral Hazard in Teams. *The Bell Journal of Economics*, 13 (2):324–340, 1982.

Holmstrom, B. and Milgrom, P. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics & Organization*, 7:24, 1991.

Holt, C. A. and Laury, S. K. Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655, 2002.

Johnson, D. W. and Johnson, R. T. *Cooperation and competition*. Theory and research, 1989.

Kachelmeier, S. J., Reichert, B. E., and Williamson, M. G. Measuring and motivating quantity, creativity, or both. *Journal of Accounting Research*, 46(2):341–373, 2008.

Kandel, E. and Lazear, E. P. Peer pressure and partnerships. *Journal of political Economy*, 100(4):801, 1992.

Kerlinger, F. N. and Lee, H. B. *Foundations of Behavioral Research*. 2000.

Kilduff, G. J., Elfenbein, H. A., and Staw, B. M. The psychology of rivalry: A relationally dependent analysis of competition. *Academy of Management Journal*, 53(5):943–969, 2010.

Kluger, A. N. and Denisi, A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2):254–284, 1996.

Koenker, R. *Quantile Regression*. Cambridge, 2005.

Koenker, R. and Bassett, G. Regression quantiles, in: Econometrica. *Journal of the Econometric Society*, pages 33–50, 1978.

Ku, G., Malhotra, D., and Murnighan, J. K. Towards a competitive arousal model of decision-making: A study of auction fever in live and Internet auctions. *Organizational Behavior and Human Decision Processes*, 96(2): 89–103, 2005.

Lazear, E. P. and Rosen, S. Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89(5):841–864, 1981.

Leary, M. R. and Tangney, J. P. *Handbook of Self and Identity*. 2005.

Levene, H. Robust tests for equality of variances1. In *Contributions to probability and statistics*, volume 2, pages 278–292. Essays in honor of Harold Hotelling, 1960.

Libby, R., Bloomfield, R., and Mark, M. W. Experimental research in financial accounting. *Accounting, Organizations and Society*, 27(8):775–810, 2002.

Locke, E. A. and Latham, G. P. *A theory of goal setting & task performance*. New Jearsy, 1990.

Lockwood, P., Jordan, C. H., and Kunda, Z. Motivation by positive or negative role models: regulatory focus determines who will best inspire us. *Journal of personality and social psychology*, 83(4):854, 2002.

Loewenstein, G. Experimental economics from the vantage-point of behavioural economics. *The Economic Journal*, 109(453):25–34, 1999.

Lynn, M. and Snyder, C. R. *Uniqueness seeking, in: Handbook of positive psychology*. Charles R. Snyder and Shane J. Lopez, Oxford, 2002. 395-410.

Malhotra, D. The desire to win: The effects of competitive arousal on motivation and behavior. *Organizational Behavior and Human Decision Processes*, 111(2):139–146, 2010.

Mann, H. B. and Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, pages 50–60, 1947.

Marquaridt, D. W. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12(3):591–612, 1970.

May, M. A. and Doob, L. W. Competition and cooperation. *Social Science Research Council Bulletin*, 1937.

Mead, M. *Cooperation and competition among primitive peoples*. New York, 1937.

Meder, M., Plumbaum, T., and Hopfgartner, F. Perceived and actual role of gamification principles. In *Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*, pages 488–493, 2013.

Moldovanu, B. and amd X. Shi, A. S. Contests for Status. *Journal of Political Economy*, 115(2):338–363, 2007.

Morey, N. and Gerber, G. L. Two Types of Competitiveness: Their Impact on the Perceived Interpersonal Attractiveness of Women and Men. *Journal of Applied Social Psychology*, 25(3):210–222, 1995.

Murthy, U. S. The Effect of Relative Performance Information under Different Incentive Schemes on Performance in a Production Task. Working Paper, 2011.

Mussweiler, T. 'Seek and ye shall find': ntecedents of assimilation and contrast in social comparison. *European Journal of Social Psychology*, 31(5): 499–509, 2001.

Mussweiler, T. Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, 110(3):472–489, 2003.

Nalebuff, B. J. and Stiglitz, J. E. Prizes and incentives: towards a general theory of compensation and competition. *The Bell Journal of Economics*, pages 21–43, 1983.

Niederle, M. and Vesterlund, L. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics*, 122:1067–1101, 2007.

Nordstrom, R., Lorenzi, P., and Hall, R. V. A review of public posting of performance feedback in work settings. *Journal of Organizational Behavior Management*, 11(2):101–124, 1991.

Payne, J. W., Bettman, J. R., and Johnson, E. J. *The adaptive decision maker*. Cambridge, 1993.

Poortvliet, P. M., Janssen, O., Yperen, N. W. V., and de Vliert, E. V. Low ranks make the difference: How achievement goals and ranking information affect cooperation intentions. *Journal of Experimental Social Psychology*, 45(5):1144–1147, 2009.

Prendergast, C. The provision of incentives in firms. *Journal of economic* *literature*, pages 7–63, 1999.

Reeves, B. and Read, J. L. *Total Engagement: How Games and Virtual Worlds Are Changing the Way People Work and Businesses Compete*. Boston, 2013.

Ryckman, R. M., Hammer, M., Kaczor, L. M., and Gold, J. A. Construction of a Hypercompetitive Attitude Scale. *Journal of Personality Assessment*, 55 (3-4):630–639, 1990.

Ryckman, R. M., Hammer, M., Kczor, L. M., and Gold, J. A. Construction of a personal development competitive attitude scale. *Journal Of Personality Assessment*, 66(2):374–385, 1996.

Rydval, O. and Ortmann, A. How financial incentives and cognitive abilities affect task performance in laboratory settings: An illustration. *Economics Letters*, 85(3):315–320, 2004.

Schmidt, K. M. Social Preferences and Competition. *Journal of Money, Credit and Banking*, 43:207–231, 2011.

Schnell, R., Hill, P. B., and Esser, E. *Methoden der empirischen Sozialforschung*. 2008.

Schreck, P. *Honesty in managerial reporting: How competition affects the benefits and costs of lying, in: Critical Perspectives on Accounting*. 2014.

Sherman, D. K. and Cohen, G. L. The psychology of self-defense: Self-affirmation theory, in: Advances in experimental social psychology. *Advances in experimental social psychology*, pages 183–242, 2006.

Silverman, R. E. Latest Game Theory: Mixing Work and Play, The Wall Street Journal. The Wall Street Journal, 2011. URL http://online.wsj.com/news/articles/SB10001424052970204294504576615371783795248. Downloaded: 27.04.2014.

Simon, H. A. Behavioural economics. *A dictionary of economics*, pages 221–224, 1987.

Simsek, Ö. F. and Yalicetin, B. I feel unique, therefore I am: The development and preliminary validation of the personal sense of uniqueness (PSU) scale. *Personality and Individual Differences*, 49(6):576–581, 2010.

Smith, V. L. Economics in the Laboratory. *The Journal of Economic Perspectives*, pages 113–131, 1994.

Snyder, C. R. and Fromkin, H. L. *Uniqueness: The human pursuit of difference*. New York, 1980.

Sprinkle, G. B. Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society*, 28(2):287–318, 2003.

Stapel, D. A. and Koomen, W. Competition, cooperation, and the effects of others on me. *Journal of personality and social psychology*, 88(6):1029, 2005.

Stapel, D. A. and Schwinghammer, S. A. Defensive social comparisons and the constraints of reality. *Social Cognition*, 22(1):147–167, 2004.

Steele, C. M. The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in experimental social psychology*, 21(2):261–302, 1988.

Suls, J. E. and Wheeler, L. E. *Handbook of social comparison*. Theory and research, New York, 2000.

Suls, J. E., Martin, R., and Wheeler, L. Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, 11(5): 159–163, 2002.

Tafkov, I. D. Private and Public Relative Performance Information under Different Compensation Contracts. *The Accounting Review*, 88(1):327–350, 2013.

Tassi, F. and Schneider, B. H. Task-Oriented Versus Other-Referenced Competition: Differential Implications for Children's Peer Relations. *Journal of Applied Social Psychology*, 27(17):1557–11580, 1997.

Taylor, S. E., Wood, J. V., and Lichtmann, R. R. It could be worse: Selective evaluation as a response to victimization. *Journal of Social Issues*, 39(2): 19–40, 1983.

Tesser, A. Toward a self-evaluation maintenance model of social behavior. *Advances in experimental social psychology*, 21:181–227, 1988.

Tesser, A., amd J. C. Collins, N. C., Cornelll, D., and Beach, S. Confluence of self-esteem regulation mechanisms: On integrating the self-zoo. *Personality and Social Psychology Bulletin*, 26(12):1476–1489, 2000.

The Economist. "Gamification", 2012a. URL http://www.economist.com/news/21566581-gamification.

The Economist. More than just a game, 2012b. URL http://www.economist.com/news/business/21565926-video-games-are-behind-latest-fad-management-more-just-game.

TopCoder, 2014. URL http://www.topcoder.com/. Downloaded: 01.05.2014.

Tor, A., Garcia, S. M., and Schiff, T. M. *The Psychology of Competition: A Social Comparison Perspective*. 2013.

Tran, A. and Zeckhauser, R. Rank as an inherent incentive: Evidence from a field experiment. *Journal of Public Economics*, 96(9-10):645–650, 2012.

Triplett, N. The Dynamogenic Factors in Pacemaking and Competition. *The American Journal of Psychology*, 9(4):507–533, 1898.

v. Dijk, F., Sonnemans, J., and v. Winden, F. Incentive systems in a real effort experiment. *European Economic Review*, 45(2):187–214, 2001.

Vorberg, D. and Blankenberger, S. Die Auswahl statistischer Tests und Maße. *Psychologische Rundschau*, 50(3):157–164, 1999.

Wald, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482, 1943.

Werbach, K. and Hunter, D. *For the Win: How Game Thinking Can Revolutionize Your Business*. Philadelphia, 2012.

Wheeler, L. Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, 1:27–31, 1966.

White, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Journal of the Econometric Society*, pages 817–838, 1980.

Wills, T. A. Downward comparison principles in social psychology. *Psychological bulletin*, 90(2):245, 1981.

Wood, J. V. Theory and research concerning social comparisons of personal attributes. *Psychological bulletin*, 106(2):231, 1989.

Wood, R. E. and Locke, E. A. Goal setting and strategy effects on complex tasks. *Research in Organizational Behavior*, 12:73, 1990.