

When the Meaning of Work Has Disappeared: Experimental Evidence on Employees' Performance and Emotions

Adrian Chadi,^a Sabrina Jeworrek,^a Vanessa Mertins^b

^aInstitute for Labour Law and Industrial Relations in the European Union, Trier University, 54296 Trier, Germany; ^bUniversity of Vechta, 49377 Vechta, Germany

Contact: chadi@iaaeu.de (AC); jeworrek@iaaeu.de (SJ); vanessa.mertins@uni-vechta.de (VM)

Abstract. This experiment tests for a causal relationship between the meaning of work and employees' motivation to perform well. The study builds on an existing employer–employee relationship, adding realism to the ongoing research of task meaning. Owing to an unexpected project cancellation, we are able to study how varying the information provided about the meaning of previously conducted work—without the use of deception, but still maintaining a high level of control—affects subsequent performance. We observe a strong decline in exerted effort when we inform workers about the meaninglessness of a job already done. Our data also suggests that providing a supplemental alternative meaning perfectly compensates for this negative performance effect. Individual characteristics such as reciprocal inclinations and trust prompt different reactions. The data also show that the meaning of work affects workers' emotions, but we cannot establish a clear relationship between emotional responses and performance.

Keywords: meaning of work • worker motivation • real-effort task • field experiment • reciprocity

1. Introduction

Information sharing is a key topic in management practice and research. Open, transparent communication across the organization is generally associated with improved employee motivation and organizational outcomes. A lack of communication might demonstrate insufficient confidence, possibly leading to distrust (Searle et al. 2011). Whereas informing employees about favorable organizational developments clearly has a positive impact on overall work attitudes, the more difficult question for a supervisor is how to deal with rather inconvenient information. Informing employees about the cancellation of a project after significant investments of time and effort might lead to negative effects on subsequent work performance. Imagine, for instance, a software development team that finds out that their painstakingly written code will be made obsolete by an operating system; or a research assistant who works feverishly on a manuscript and then finds out that it was never submitted. Virtually nothing is known about the consequences of information sharing when work that has already been performed loses its meaning. The present paper is the first to shed light on the question of how the communicated meaning of previously completed work affects workers' current performance.

There is a long research tradition suggesting that employees' attitudes and workplace behavior are

sensitive to whether they perceive their assigned tasks to be meaningful.¹ By contrast, there is relatively little research that attempts to establish a clear causal link between task meaning and work performance. A majority of studies relies on cross-sectional designs, failing to rule out the possibility that task meaning is a consequence and not a cause of job performance. Some recent experiments in behavioral economics (Ariely et al. 2008, Bäker and Mechtel 2013, Chandler and Kapelner 2013) are able to isolate this effect from potentially confounding factors.

Experimental investigations are often considered a major source of knowledge in the social sciences (Falk and Heckman 2009). Testing the effects of the meaning of work within experimental settings, however, may be subject to limitations. Recent discussions on generalizability raise questions relating to the extent to which results from laboratory experiments extrapolate to real workplace behavior (Levitt and List 2007, Al-Ubaydli and List 2015). Aspects such as the naturalness of employer–employee relationships and the context in which work experience is embedded may be of concern. When subjects know that they are taking part in an experiment, and therefore make decisions in more or less artificial settings, they may also form beliefs about the experimental objectives and associated appropriate behavior (Zizzo 2010). Regarding

the overall quest for generalizability, the characteristics of previous experiments on task meaning suggest that this research field could benefit from taking further steps toward creating a more natural workplace scenario.

So far, both laboratory experiments (e.g., Ariely et al. 2008, Bäker and Mechtel 2013) and field experiments (e.g., Grant 2008, Chandler and Kapelner 2013, Kosfeld et al. 2017) consider specific manifestations of task meaning, telling participants in “no meaning” conditions, for instance, that their future work is futile a priori, that no use would ever be made of its outcome, or even that it would be destroyed shortly after completion. These setups are not only helpful in identifying underlying mechanisms or channels behind the meaning-in-labor effects; they are also useful in understanding workers’ behavior in the context of minimal meaning and Sisyphus’ labor. Frequently, workers anticipate the possibility that their work might never be used, so that assigned work tasks are perceived as unrewarding and pointless. Nevertheless, employment relations are of course primarily intended for the production of marketable or otherwise valuable outcomes, so that workers can typically expect their work to be worthwhile. If their actual work turns out to be meaningless, we argue that this is very often a retrospective experience. Supervisors assign tasks, workers exert a lot of effort, and then decision makers cancel the project or the work becomes futile for some other reason. Empirical research on such occurrences, as commonplace and detrimental for organizational outcomes as they probably are, is not available within the previous literature.

The present study tries to shed light on the existence and determinants of a causal relationship between the meaning of previously completed work and workers’ current performance. A proper research design would consist of a natural environment, involving an authentic employer–employee relationship, in which workers are committed to do a useful job while the researcher maintains control over the decision context. At the same time, it must be possible to credibly guide part of the workforce to believe in a sudden loss of meaning. Such restrictive conditions can best be met by means of a controlled experiment containing field characteristics (Harrison and List 2004), and since experimental economics proscribes the use of deception for several well-founded reasons (see, e.g., Jamison et al. 2008), the loss of a job’s meaning would ideally occur as a matter of fact.

Against this background, the present study exploits the particular happenstance of having access to a number of temporary workers that were part of a recently conducted field experiment (Jeworrek and Mertins 2014). For these employees, the meaning of the work task (setting up a database on business reports for

research purposes) did indeed subsequently disappear. This unexpected turn of events allows us to study—without the use of deception, but still maintaining a high level of control—how simple variation of the information available on the meaning of previously completed work affects workers’ current performance. We achieve this variation by short text modules that reveal all, part, or none of the information about the outcome of the project.

Accordingly, we randomly assigned each worker to one of three information sets. In one group, we simply reminded workers about the meaning of their previous task (i.e., creating a database for research purposes), so that they were likely to still believe in the ORIGINAL MEANING. Additionally, we told the NO MEANING group that the managing director had decided to dissolve the archive after the creation of the database. To study how employees respond to a new purpose after the original one had disappeared, a third group received an ALTERNATIVE MEANING treatment: We informed employees that their previous work had become meaningless with regard to its stated purpose, but we also told them that their work served an additional purpose not previously communicated (in this case, research in the field of personnel economics). Afterward, we measured and compared employees’ willingness to exert real effort across the three groups. To isolate the pure effect of task meaning from the interplay with monetary incentives, we implemented an incomplete contract with a fixed payment independent of performance. Under these conditions, standard neoclassical theory suggests that rational payoff-maximizing individuals will exert no effort at all, and information about the meaning of previously completed work does not alter this prediction. Recent empirical evidence, however, suggests that people care about the meaning of their tasks, although reported effects were often rather weak. This is what we expect here as well, especially as we put an emphasis on manipulating the perceived task meaning in a natural and respectful way. In addition, the focus here is on the meaning of a job already done, and this retrospective approach, with a considerable time lag between the tasks, might be expected to weaken employees’ responses.

Despite this, the data reveal a strong impact of previous task meaning on workers’ subsequent motivation. Furthermore, employees seem to “forgive” their employer for canceling a project if another and still valid purpose is credibly communicated. The treatment effects are robust across various specifications but are driven by those workers who answered our survey invitation within a few days. A particular subgroup of employees that we call latecomers did not respond to the survey invitation until they received a final call. These workers do not seem to be susceptible at all to the variation in task meaning.

The study also analyzes the role of emotional reactions in moderating employees' responses to the meaning of labor. Until now, the idea that negative emotions may translate into decreased performance has been theoretically assumed (Compte and Postlewaite 2004) but not empirically demonstrated.² Here, the natural setting enables an analysis of the role of emotions in the workplace. Our analysis suggests that revealing the loss of the previous work's meaning indeed triggers negative emotions, whereas providing the alternative meaning compensates for this negative effect. In contrast to the above assumption, however, we do not observe a clear link between negative emotions and low performance.

Finally, we extend the existing literature by exploring the possibility that not all individuals respond uniformly to the intervention. First, we explore the impact of individuals' reciprocal inclination on their willingness to respond with cooperative effort to the assignment of meaningful work, which might be considered a nonmonetary gift.³ In fact, especially highly reciprocal individuals are the ones who respond positively to the existence of task meaning. Second, we show that an individual's level of innate trust plays a crucial role in the ALTERNATIVE MEANING treatment. Distrustful individuals perform significantly worse than trustful ones, probably because they might suspect that the supplied alternative task meaning would become similarly futile.

2. Experimental Design and Data

2.1. Background

The experiment was designed to measure the effect of retrospective task meaning on performance in a subsequent unrelated task. This was done by looking at the same workers tackling two different tasks in order to establish whether *information about the meaning of Task I* (treatment variation) affects workers' *performance in Task II* (experimental measure).

The task environment was provided by a German research institute, which has collected business reports (approximately 14,000 issues) over decades without collecting adequate corresponding inventory data. In the course of the institute's realignment, the managing director initiated an inventory update for the so-called "Unternehmensarchiv." In September 2013, 140 randomly selected workers from a pool of applicants carried out the task of filing business reports (Task I). They were hired for a one-time half-day job (3.5 hours), working in isolation in their assigned offices to prevent peer effects (see, e.g., Falk and Ichino 2006). The task consisted of sequentially picking business reports (arranged in random order) from a nearby shelf and typing the relevant information (such as company name, year of publication, and number of pages) into an electronic entry mask. Jeworrek and

Mertins (2014) used this real-life working scenario to study the causal relationship between wages, workers' autonomy, and performance. Workers were randomly assigned to one of the experimental treatments, which varied according to paid wage (ranging from €30.00 to €42.00), wage determination procedure (workers either self-determined their wage or were allocated a wage randomly), and knowledge of coworkers' autonomy (awareness of whether or not others determined their wage).

Importantly, for the present study, the content of Task I was identical for all workers, but we varied information about its meaning *ex post* simply by providing or withholding updates about the extent to which the task still served its original stated purpose. At first, the collection of business reports was perceived as comprehensive, unique, and therefore a valuable resource for researchers. Subsequently, the managing director decided to dissolve the archive when the inventory revealed the collection's fragmentary character, the bad condition of some reports, and the increasing online accessibility of such reports. As a consequence, the effort previously invested in the project no longer served its original stated purpose. At the same time, workers were unaware of the second purpose of contributing to a research project, and we decided to exploit this happenstance—which from a researcher's perspective, it was—by varying the information supplied to randomly selected groups of workers in order to study their subsequent motivation to perform an additional task.

2.2. Implementation

Approximately one week after the undertaking of the inventory, we contacted workers and invited them to take part in an incentivized follow-up questionnaire (see Section B.1 in Online Appendix B). In the invitation email, we informed participants that this online survey sought to gain some scientific insights into employees' recent experiences at work. We made clear that this was the last interaction between employer and employee, so that employees' subsequent behavior was likely to reflect decision making in a one-shot situation.

On completion of the first part of the survey—collecting information about sociodemographic data, personality traits, and individuals' underlying motives regarding the research questions studied in Jeworrek and Mertins (2014)—all employees earned a flat payment of €5.00, as well as some additional money for incentivized decision making, resulting in an average payment of €10.97 for a mean processing time of 42 minutes in total (including Task II). After finishing the first part of the survey, we emphasized that any effort exerted on the second part of the questionnaire (Task II) would support ongoing research activities, but that earnings were now fixed and did not depend

on workers' subsequent behavior (see Section B.2 in Online Appendix B). This situation mirrors the incomplete contracts that are typical of employment relationships. The measure of effort (described in more detail below) is based on participants' responses to a series of similar survey questions or vignettes.

After clarifying the details of payment, we implemented three treatment groups by randomly allocating each worker to one of three information sets, which differed only in respect to the information provided about the meaning of Task I. Following the two-component conceptualization of meaning in labor by Ariely et al. (2008), we varied the purpose of the task while keeping recognition constant by ensuring that we, first of all, acknowledged employees' completed work. The ORIGINAL MEANING group was simply reminded of the previously communicated purpose of their work task (i.e., creating a useful database of business reports). In the ALTERNATIVE MEANING treatment group, we initially informed employees that their work was meaningless with regard to its original purpose but afterward, we disclosed that their previous work served another purpose (i.e., contributing to research in the field of personnel economics). In this way, we renewed the meaning of the previous task. By revealing only the loss of original meaning to the third treatment group, NO MEANING, we implemented a comparison group of randomly selected individuals who would no longer be expected to find any task meaning in their previous work. The exact wording of the treatment variations is set out below:

[All participants]

We would like to provide you with some additional information concerning your work for the business archive. As you have seen, the yearly reports from various German and international companies have been collected over the years. There has, however, never been a systematic archiving of the data. Your task, and the task of the other hired data entry assistants, was to enter the yearly reports into a data bank, where they can be used for research purposes.

[Additional for NO MEANING and ALTERNATIVE MEANING]

For various reasons such as the increasing online accessibility of company data, we assume in the meantime that the data you have processed will be of little scientific interest and that your work was, against this background, to a great extent futile.

[Additional for ALTERNATIVE MEANING]

By contrast, personnel economics research is very interested in data that represents real work scenarios. As a result of your employment, the prerequisites for a number of meaningful scientific investigations are now available.⁴

If workers within the NO MEANING group subsequently perform worse than workers in the ORIGINAL/

ALTERNATIVE MEANING treatments, at least two different interpretations are conceivable. First, information about the meaning of past work may influence expectations about the meaningfulness of the new work, which in turn affects performance in the second task [*spillover effect*]. Although we clearly stated the meaningfulness and importance of the second task in the questionnaire (see Section B.2 in Online Appendix B, General Instructions), those individuals who received additional information (NO MEANING/ALTERNATIVE MEANING) might have updated their expectations on the likelihood that the new work will be rather useful or a waste of time. If this is a valid explanation, then giving an alternative meaning to more trustful workers should result in a higher performance than giving the same information to less trusting people. We examine this interpretation further in the course of the empirical analysis (see Section 3.2).

Second, the additional information may trigger emotions that in turn cause behavioral responses [*mood effect*]. In particular, workers who have learned about the archive's dissolution might feel angry or may be disappointed that their time was to some extent wasted. Receiving an alternative meaning might dampen such feelings. Although the business reports filed by workers have been shredded in the end (see the pictures in Section B.5 of Online Appendix B), we decided to withhold this information in favor of a more natural communication strategy. It follows that the resulting treatment effects are expected to be at the weaker end of the potential NO MEANING response range. To test whether or not revealing such information in fact induces negative feelings, we measured workers' emotions shortly after the treatment, using an adapted version of the Positive Affect Negative Affect Scale (PANAS) by Watson et al. (1988), a tool frequently used by organizational psychologists (for a German version, see Krohne et al. 1996). Participants indicated on a 5-point scale how they were feeling at that moment (ranging from 1 = "very slightly or not at all" to 5 = "very much").⁵ Because of the study's focus on workers' potentially negative emotional reactions to being informed about the meaninglessness of previous work, only differences regarding the emotions *upset*, *disappointed*, *replaceable*, and *shocked* are of interest, as we expect workers in the NO MEANING treatment to score higher on these items than individuals in the other groups. We mixed these "target emotions" with 20 completely unrelated emotions (fillers) such as *active*, *strong*, or *attentive* for two reasons: First, their use should distract participants from the purpose of this part of the questionnaire and help to minimize experimenter demand effects (Zizzo 2010). Second, the fillers allow checking for whether the application of randomization into treatments was successful by testing for

differences in general sentiment between the treatment groups.

Afterward, we conducted a real-effort experiment in which workers were, however, not told that they were participating in an experiment. We asked all participants to judge various standard workplace situations (see Sections B.2 and B.4 in Online Appendix B for more information) as fully as possible, but we allowed them to skip questions or even to quit the survey at any time. Because participants received a fixed fee irrespective of their response behavior, they had an incentive to shirk by quitting the questionnaire at an early stage, or by not working on it at all. Accordingly, participants' performance in the vignette study can be interpreted as a reciprocal gift (Akerlof 1982; Akerlof and Yellen 1990; Fehr et al. 1993, 1998; Gächter and Fehr 2002) in exchange for monetary and nonmonetary gift giving by the employer. In this situation, it is an open empirical question as to whether and how varying the meaning of work may affect participants' perceptions and subsequent behavior.

We presented the workplace situations as vignettes, implying only slight changes in the depicted setting. Workers needed to provide a total of 48 answers to not be classified as ending the survey early. Workers did not know the large number of vignettes *ex ante*. Responding to vignettes involved opportunity costs for all participants because this is usually considered to be a monotonous task. We use the number of answered vignettes as the behavioral measure of participants' willingness to exert effort. This is similar to a recent study by Becker et al. (2013) who interpret survey response behavior as an economic outcome measure.

Because the study is built on recently employed workers performing the same task in a temporary job, the population of participants was by design restricted to 140 potential candidates. Far more than half of the workers participated immediately upon receiving an invitation, but we had to exclude some individuals who had finished the survey before the treatment text was shown from the analysis, leaving a total of 81 participants ($n = 25$ in ORIGINAL MEANING; $n = 27$ in NO MEANING; $n = 29$ in ALTERNATIVE MEANING). To analyze the treatment effects among the remaining former workers too, we aimed at motivating as many non-respondents as possible to participate. The sending out of further invitation emails during the following weeks attracted 26 additional participants of which 21 were confronted with the treatment texts. Therefore, we ended up with a pooled sample of 102 respondents, among which, however, we consider the latecomers to be a special group. The potentially self-chosen delay in participation suggests that they may differ from the core group in individual characteristics as well as in their behavioral reactions to the experimental intervention. Table 1 summarizes the sequence of events.

Table 1. Timeline of the Experiment

Activity	Description
1. Task I	Filing of business reports (with treatment variation)
2. Invitation	Invitation for survey participation via email
3a. First part of the survey	Questionnaire on sociodemographics, personality, preferences, and issues related to Task I
3b. Payment	Payment is fixed by now
3c. Second part of the survey	Description of Task II
3d. Experimental treatment	Varying levels of information on meaning of Task I
3e. PANAS	Questionnaire on emotions
3f. Task II	Identical real-effort task with 16 vignettes and 3 questions each (output $\in [0; 48]$)

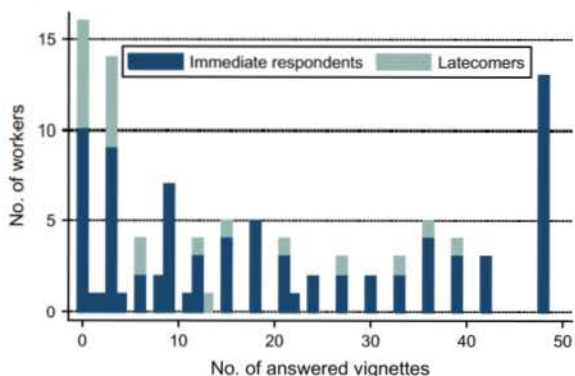
In step 3a, we gathered data on workers' sociodemographics, personality, and preferences, which may help to explain the heterogeneity in employees' performance. In particular, we administered a 40-item (German language) version (Weller and Matiaske 2009) of the well-established NEO Five-Factor Inventory by Costa and McCrae (1989) to measure individuals' Big Five personality traits.⁶ We also introduced a simple incentivized trust game (Berg et al. 1995) with participants playing the roles of both sender and receiver. The amount sent to the receiver is a measure of trust. In the role of receiver, the amount sent back to the sender measures trustworthiness or reciprocal inclination.

3. Analysis

3.1. Performance

As seen in Figure 1, the performance of the workers, measured by vignette responses during the real-effort task, shows considerable variation: Although only 16%

Figure 1. (Color online) Overall Performance in Vignette Task



of the workers refused to complete any vignette (pursuing the rational strategy in the absence of monetary incentives), a clear majority was at least induced to work. A small group (13%) of former workers responded to the whole set of vignettes. On average, workers answered 18.32 vignettes out of a maximum of 48. As indicated by Figure 1, response rates were lower among the above-mentioned latecomers (light blue bars), compared to those who immediately took part in the survey (dark blue bars). The immediate respondents completed on average 20.32 vignettes.

Figure 2 demonstrates that the average number of responses differs between the three treatment groups. We start by discussing the results for the pooled sample (left bars). Later on, we focus on immediate survey respondents only (right bars), which excludes the above-mentioned 21 survey latecomers from the analysis because of their potentially differing behavior.

Compared with the No MEANING group (mean = 13.88), the average number of provided responses is higher among both the ORIGINAL MEANING group that was not informed about the loss of the original meaning from the previous job (mean = 18.47), and the ALTERNATIVE MEANING group that was briefed about the alternative meaning (mean = 22.37) (for additional descriptive results, see Section A.2 in Online Appendix A). As indicated by the large confidence intervals in Figure 2, the variation in response behavior is high, and so these group differences, though obvious, are not necessarily expressed in p -values of statistical significance tests. Within the pooled sample, performance is significantly higher in the ALTERNATIVE MEANING group than in the No MEANING group with $p = 0.055$, whereas the performance difference between the ORIGINAL MEANING group and the No MEANING group is not statistically significant (one-sided Wilcoxon rank-sum tests). The inclusion of latecomers seems to dilute the treatment effects: when focusing on immediate survey respondents, we observe statistically significant

Figure 2. (Color online) Treatment Effects

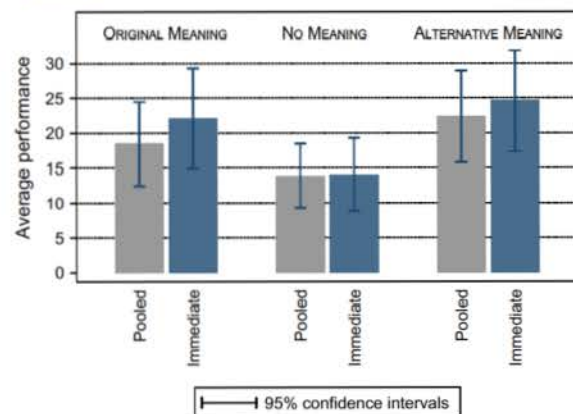


Table 2. Main Regression Results

	(1)	(2)	(3)	(4)
No MEANING	Reference category			
ORIGINAL MEANING	4.592 (3.713)	7.059* (3.897)	8.242** (3.781)	11.486** (4.541)
ALTERNATIVE MEANING	8.493** (3.939)	11.426*** (3.938)	11.363*** (3.819)	15.022*** (4.427)
Latecomer			-9.224** (4.555)	
ORIGINAL MEANING × Latecomer				-13.600* (6.923)
No MEANING × Latecomer				2.532 (6.491)
ALTERNATIVE MEANING × Latecomer				-14.650* (7.816)
Constant	13.879*** (2.545)	14.744*** (4.170)	16.466*** (4.226)	13.549*** (4.617)
Controls	No	Yes	Yes	Yes
N	102	102	102	102
Adjusted R ²	0.023	0.101	0.141	0.157

Notes. The dependent variable is the number of vignette responses. Robust standard errors are in parentheses. All linear variables are demeaned. Controls include sociodemographics, the Big Five, and organizational controls.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

differences for both treatments compared with No MEANING with $p = 0.056$ and $p = 0.024$, respectively.

Therefore, we study the differences between immediate respondents and latecomers who answered only after receiving further invitation emails by including a latecomer dummy variable within the regression analysis. Furthermore, we check the role of differences in workers' characteristics by controlling for several individual factors using ordinary least squares (OLS).⁷

The estimation results in Table 2 illustrate that the additional controls strengthen the observed performance effects with respect to the size of the estimated effects as well as the corresponding significance levels. We control for the basic sociodemographics *age*, *gender*, and *nationality*, and for the Big Five personality traits since (i) previous research suggests a link between individual performance and personality in the context at hand (Grant 2008) and (ii) the present data indicate that certain personality traits are not equally distributed across treatments (see Section A.3 in Online Appendix A). We also add organizational controls for the time of the day when respondents completed the questionnaire (*morning*, *noon*, *afternoon*, and *evening*) since the time may have affected response behaviour (see also Section A.4 in Online Appendix A for the complete estimation results). Including these controls in specification (2) considerably increases the explanatory power of the model, and the positive performance effects of both meaning treatments appear to be significant at the 10% and the 1% level, respectively. When comparing the treatment averages as well

as the estimated coefficients for both meaning treatments, it becomes obvious that providing an alternative meaning for the previously conducted job absorbs the negative performance effect of the loss of the original meaning. Apparently, it does not matter that the task's original meaning was lost: what counts is that it had a meaning at all. The following specification (3) additionally accounts for the fact that the sample consists of two distinct groups, immediate respondents and latecomers. It shows that these latecomers exert much less effort by answering approximately nine vignettes less than the immediate respondents. This is a sizeable amount compared with the average number of 20.32 vignettes answered by immediate respondents. A two-sided Wilcoxon rank-sum test confirms that, across all the treatments, the behavior of immediate respondents and latecomers differs significantly with $p = 0.013$ (for further descriptive results and nonparametric statistics, see Section A.5 in Online Appendix A).

The group of latecomers is peculiar for at least two reasons: First, more time had already elapsed by the time these latecomers had answered the vignettes, probably going along with a less intense retrospection about the former job so that the effects of task meaning relating to work already done simply faded away.⁸ Second, latecomers had chosen not to answer the questionnaire upon the first request. Although there might be exogenous reasons for this, such as limited reachability during holidays, it could also constitute a self-selection related to particular personal characteristics. For instance, one may think of low loyalty toward the employer or low interest in continuing the relationship. In any case, both lines of arguments suggest heterogeneous effects between immediate respondents and latecomers, since the latter are expected to be less motivated than the core group to answer vignettes across treatments and presumably less responsive to the information about the meaning of the task.

As mentioned before, we observe significant treatment effects within the core group of immediate respondents. According to nonparametric test statistics, latecomers, however, do not respond at all to the different information sets. Specification (4) in Table 2, which includes interaction dummies between the three treatments and latecomers, supports this finding. The interaction terms for the treatments ORIGINAL MEANING and ALTERNATIVE MEANING are significantly negative in sign and approximately as large as the estimated treatment coefficients. In the remaining Sections 3.2 and 3.3, where we try to shed some light on the underlying mechanisms that have caused the observed meaning effects, we focus on the immediate respondents since our previous analysis has shown that latecomers are not affected at all by varying information sets.

Similar to the high variance observed in the number of answered vignettes, there is an even larger variation in the number of words written by respondents, ranging from 0 to 738 with a resulting average of 128.29 words and a standard deviation of 137.44. Since an individual's effort can be expressed in terms of the number of answered vignettes as well as of the number of words written, we run specification (3) again with the number of words as dependent variable to check the robustness of our findings within the pooled sample. The estimated coefficient for ORIGINAL MEANING is 49.59 ($p = 0.081$), and for ALTERNATIVE MEANING it is 65.78 ($p = 0.088$). Therefore, we observe significant meaning effects in both treatments, with a resulting pattern that is very similar to the previous results. However, because respondents were asked to answer as many vignettes as possible, we adopt the total number of answered vignettes for the remainder of the analysis as a more appropriate measure of individual performance.⁹

Because this experiment on the role of retrospective task meaning is embedded in a former employer-employee relationship studied in a field experiment, we also consider aspects of the previous job in checking the robustness of the results. The findings remain robust when we include diverse additional variables reflecting the previous work relation such as *wage earned before*, *previous treatments*, and *previous performance*. Differences in earnings of up to €12.00 and variation in previous experimental treatments have no influence on the main findings. Similarly, workers' ability to perform the previous task is clearly unrelated to performance in the present task, confirming that both tasks are indeed different (for complete results, see Section A.6 in Online Appendix A). Taken together, these facts indicate the robust effects on performance induced by the information supplied about the previous work task.

Result 1. The meaning of a previous task significantly influences workers' performance, even in an unrelated follow-up task. A loss of meaning results in a considerable drop in performance. This drop, however, is completely compensated for when an alternative meaning for the previously conducted job is provided.

3.2. Emotions

In this subsection, we investigate the suggestion that a loss of meaning also influences respondents' emotional states, and that these in turn might affect workers' performance. When we compare the average scores¹⁰ for the four selected emotions (see Table 3), we find that the No MEANING treatment triggered negative emotional reactions: respondents in this group felt significantly more *disappointed*, *replaceable*, *shocked*, and *upset*.¹¹ Note that all negative emotions are at generally low levels. This is not surprising since perceptions

Table 3. Emotions

	ORIGINAL MEANING	NO MEANING	ALTERNATIVE MEANING	ORIGINAL VS. NO MEANING	ALTERNATIVE VS. NO MEANING
	Means			<i>p</i> -values	
Disappointed	1.44	1.85	1.21	0.075*	0.008***
Replaceable	1.44	1.92	1.57	0.010**	0.040**
Shocked	1.36	1.80	1.26	0.023**	0.012**
Upset	1.28	1.65	1.32	0.035**	0.032**

Note. (One-sided) Wilcoxon rank-sum tests are carried out using the ordinal variables with five categories.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

of overall working conditions are very good (see Section A.3 in Online Appendix A). It does not seem to matter whether we informed workers about the ALTERNATIVE MEANING or just reminded them of the ORIGINAL MEANING: the average emotional state is almost identical across the two groups.¹² Like the observed average performance in the vignette task, participants' emotional reactions support the view that the alternative story neutralizes any negative effects aroused by the former task's loss of original meaning.

Although treatments obviously affect emotions, the latter do not seem to mediate the relationship between treatments and performance: if emotions were the driving force, treatment effects should not be observable for emotionally unaffected workers (i.e., workers who indicated 1 on a 5-point Likert scale). Even among such workers we do, however, observe significant differences across treatments (see Section A.8 in Online Appendix A).

Result 2. The loss of meaning triggers negative emotional reactions. These effects can be neutralized by providing an alternative meaning. Emotions, however, do not mediate the relationship between treatments and performance.

3.3. Heterogeneity

This section deals with the possibility that not all individuals respond uniformly to treatments. We consider individuals' reciprocal inclinations and their level of trust in other people as characteristics that may influence workers' reactions to the given information.

Since the task of answering vignettes was rewarded by a fixed payment, irrespective of respondents' effort in the vignette task itself, performance can be interpreted as a reciprocal action toward the employer. This invites a closer look at respondents' reciprocal inclinations. If meaning is attached to a work task, this can be seen as a nonmonetary gift, such as recognition from the employer or good working conditions, and employees may reciprocate such a gift by exerting higher levels of effort (positive reciprocity). In the NO MEANING treatment, workers might feel an obligation to pay the employer back for assigning them pointless work to

do (negative reciprocity). For these reasons, we expect highly reciprocal workers to be especially sensitive to issues of meaning in their previous job.

To check the role of reciprocity in our context, we estimate separate regressions among individuals with high and low reciprocal inclination according to a median split. Table 4 presents the regression results for both groups. We find significantly positive meaning effects ($p = 0.078$ and $p = 0.067$, respectively) only among the highly reciprocal workers. We control for the standard sociodemographics and add the previously earned wage and dummy variables for the main treatments from the former experiment. We consider the latter to be potential sources of workers' reciprocal behavior in the vignette task and therefore as relevant control variables. Even for highly reciprocal individuals, however, there is no significant influence from the previous job other than its meaning, again

Table 4. Reciprocity and Task Meaning

	(1) Low reciprocal inclination	(2) High reciprocal inclination
NO MEANING	Reference category	
ORIGINAL MEANING	1.707 (6.293)	11.73* (6.422)
ALTERNATIVE MEANING	7.694 (6.182)	12.48* (6.540)
Choice treatment	4.093 (6.459)	-1.724 (6.814)
No choice treatment	-7.921 (5.492)	-3.164 (7.575)
Wage earned before	0.963 (0.781)	0.655 (0.671)
Constant	20.63*** (6.021)	14.65** (5.711)
N	44	37
Adjusted R ²	0.198	0.113

Notes. The dependent variable is the number of vignette responses. Robust standard errors in parentheses. All linear variables are demeaned. Because of the low number of observations, we take the reduced model, including only the sociodemographic characteristics as control variables.

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

highlighting the important role of task meaning in employment relations. Furthermore, the results suggest that both mechanisms, positive and negative reciprocity, could be at play in this context.

Finally, we consider senders' behavior in an incentivized trust game since, as argued before, we cannot exclude the possibility that some people distrusted the information that the second task was meaningful. If this is a valid concern, the observed positive effect of providing an alternative meaning should be at the lower end of a universal phenomenon. To test this assertion, we differentiate the sample according to participants' general level of trust. Those who trust less are expected to expend less effort than those who trust more when informed about the alternative meaning, whereas there should be no observable difference between low and high trusting individuals in the remaining treatments. Indeed, when the sample is split in this way, there are no significant differences between individuals' performance in the ORIGINAL MEANING treatment (Wilcoxon rank-sum test, two-sided, $p = 0.577$) and in the NO MEANING treatment ($p = 0.965$). In the ALTERNATIVE MEANING treatment, however, individuals with a generally high level of trust answered 29.57 vignettes on average, whereas the less trusting answer only an average of 20 vignettes (Wilcoxon rank-sum test, one-sided, $p = 0.075$). This finding also points to a potential spillover effect: more trustful individuals seem to update their expectations on the likelihood that the second task will be meaningful in a different way than less trustful individuals. The effect is even stronger when confined to highly reciprocal individuals sensitive to task meaning (31.13 versus 10.80 vignettes, Wilcoxon rank-sum test, one-sided, $p = 0.030$). It seems that highly reciprocal individuals tend to reward their employer for a meaningful job by answering more vignettes, but only if they truly believe in the meaningfulness of their work.

Result 3. Highly reciprocal individuals and those who responded immediately drive the performance effects. A high level of trust seems to be especially important for believing in an alternative meaning and performing well afterward.

4. Conclusion

How does the perceived meaning of previously completed work affect workers' current performance? Common intuition suggests that people like to know that their effort was worthwhile, with varying perceptions of previous task meaning translating into differences in mood states and subsequent workplace behaviors. This experiment investigates the existence of such effects in a setting where real work had lost its meaning after the job was done, and where participants were paid a flat fee to perform another, unrelated

task. If people are willing to work in such a setting, then they do so voluntarily, despite the absence of explicit monetary incentives. Our results indicate that individuals are indeed willing to exert considerable effort, dependent on information about a former job's meaning: knowing that their previous work served a purpose clearly motivates individuals to do further work. It follows that employees seem to work hard not only for monetary gains or external recognition but because of an additional gain from their personal investment in the work. This extends previous causal evidence on employees' willingness to reciprocate kind behavior, shedding light on the importance of meaningful work. If employees do not care about the meaning of their work, employers could direct subordinates without any regard to such matters. Our findings, however, clearly reject this view. All things being equal, individuals perform significantly worse if they are informed that their previous work has become meaningless. This negative effect disappears if people are offered a credible alternative task meaning.

It is worthy of note that task meaning affects employees' emotions, as well as their behavior: those who subsequently find that their work was futile are significantly more disappointed, shocked, upset, and feel more replaceable than their peers. The identified treatment effects, however, do not appear to be in a clear relationship with these emotions. Our analysis also shows that certain workers respond more to meaning-in-labor than others. In particular, those who had to be convinced to participate in the survey and individuals with a low reciprocal inclination remain virtually unaffected.

These findings have some practical implications for how companies deliver feedback and organize work. Supervisors are often confronted with the dilemma of whether and how to inform subordinates that their work has turned out to be futile. The evidence suggests that providing this information in isolation is likely to have negative consequences for employees' subsequent motivation, but identifying another purpose can completely compensate for the experienced loss. Supervisors may be well advised to factor these behavioral effects into their approach when delivering such information. For example, accepting short-time marginal losses by finding another use for work already done may be beneficial in the long run when set against the potential future costs resulting from demotivated workers. If employees' experiences and perceptions from a half-day job already trigger the observed negative effects, it seems very likely that the effects on behavior and emotions arising from a loss of meaning in longer projects will be even more severe.

Finally, the analysis of heterogeneous effects has further implications. It is worth noting that highly reciprocal individuals are of central importance in such a

context. Gift exchange has proved to be a robust phenomenon in the laboratory, but empirical evidence in the field is scarce and somewhat ambiguous. In particular, it seems clear that it is not the monetary gift that induces workers to return the favor by exerting more effort (Gneezy and List 2006, Hennig-Schmidt et al. 2010), but nonmonetary gifts such as recognition (Kosfeld and Neckermann 2011), participation in decision making (Jeworrek and Mertins 2014), or the way gifts are presented (Kube et al. 2012) are known to increase performance. In identifying highly reciprocal workers as those whose performance is dependent upon the meaning of their previous job, the present study indicates that meaningfulness of work is another important source of nonmonetary gift giving.

A typical shortcoming in this research context is that the naturalness of the workplace scenario may be at the expense of sample size. This relates to an important debate on how much one should trust initial findings, which goes beyond the above discussion on laboratory versus field evidence. In a recent contribution, Maniadis et al. (2014) present a theoretical framework through which they point out several aspects that determine the probability of whether novel findings turn out to be true in the end. They advocate for repeated examinations of the same research objective to reduce the likelihood of false positives and to substantiate initial claims.

The question regarding the reliability of the findings discussed here is whether meaning plays a role in the economic context or not. Although multiple available investigations empirically support this once novel idea, as introduced by Loewenstein (1999) to economics, our investigation may serve as another piece of a puzzle indicating that meaning matters for economically relevant outcomes. Given that there are also some studies from outside the laboratory, our study conforms well to an increasing body of research. If we apply the Maniadis et al. (2014) framework rigorously to our paper, we may also discuss our findings as initial ones, since we are the first to vary the meaning of a completed task. To ensure that the declaration of a research finding is actually true, the prior probability of the hypothesis is an important aspect. With a novel research objective, research priors may per se not be very high, although it appears to us as being rather intuitive that variation in the meaning of a previous task affects workers' current motivation. Given the novelty of this type of evidence and the limited statistical power due to the small sample size, we are, however, curious about new insights through additional investigations into whether meaning indeed matters in retrospect.

This leads to further avenues for future research on task meaning, some of which we have already outlined. First and foremost, replications of our study can further

enhance the assertion that our research priors were true. Because we argue that a natural workplace scenario is important, replication attempts should ideally exploit particular happenstances, in which it is possible to credibly manipulate information on previous task meaning. To increase realism further, it would also certainly be promising to exploit settings that are characterized by long-term employer–employee relationships to investigate whether the previously identified effects of task meaning are even underestimated due to the short-term nature of the job. Such settings would also allow the analysis of the interaction of task meaning effects with various other job and workplace characteristics in permanent employment relationships. Moreover, it should be geared to a deeper understanding of the relationship between meaning of work, workers' emotions, and their performance. Even though our findings suggest that workers' emotions are not the main transmission channel, we cannot exclude the possibility that negative emotions contribute somewhat to the overall negative performance effect after a loss of meaning. Finally, our results also raise the question of exactly how the perceived meaning of the previous job affects the perception of the meaning of the current task. It seems plausible that at least some workers form expectations about the usefulness of their current work given their experiences, but future studies should find out more about this type of workplace phenomena. Laboratory experiments could be particularly helpful in identifying and isolating such spillover effects.

Acknowledgments

The authors are grateful for comments and suggestions by John List (the department editor), the associate editor, and three anonymous referees. For helpful discussions, the authors thank participants of the Workshop on the Impact of Human Resource Management in Mannheim, the Workshop on Economics in Trier, the Experimental Methods in Policy Conference in Curacao, the Annual Conference of the European Society for Population Economics in Braga, the Annual Meeting of the French Experimental Economics Association in Besancon, the Behavioral and Experimental Economics Symposium in Maastricht, the Annual Congress of the European Economic Association in Toulouse, the International Meeting on Experimental and Behavioral Social Sciences in Oxford, the European Association of Labour Economics Conference in Ljubljana, and seminar participants in Augsburg, Trier, and Vechta. The authors also thank their research assistants Manuel Hoffmann, Valentin Langholf, and Anna Stalsky for valuable support.

Endnotes

¹Research ranges from narrowly defined fields such as task significance (e.g., Hackman and Oldham 1976), where the job is said to have a substantial impact on the lives or work of other people or an organization's underlying prosocial mission (e.g., Tonin and Vlassopoulos 2010, Fehrler and Kosfeld 2014, Carpenter and Gong 2016, Ashraf et al. 2014), to a broad understanding of the meaningfulness of work

(Rosso et al. 2010, Loewenstein 1999). See also the meta-analyses by Fried and Ferris (1987) or Humphrey et al. (2007).

²See Rick and Loewenstein (2008) for a discussion of the role of emotions in economic behavior.

³Recent research suggests that a wide portfolio of nonmonetary incentives, such as recognition by the employer (e.g., Kosfeld and Neckermann 2011, Bradler et al. 2016) or simple gifts (e.g., Kube et al. 2012), plays an important role in labor relationships.

⁴We have no indications of workers' awareness of the experimental conduct. Rather, various facts argue against such a perception: the treatment phrase was framed in a studiously vague way; our stated interest in data does not necessarily imply an underlying treatment manipulation; and experimental research was virtually unknown at the place of conduct, even among economics students.

⁵Self-reported responses on 20 items (i.e., words that describe different feelings and emotions, such as strong, distressed, or guilty) are usually aggregated into two mood scales (positive and negative affect).

⁶The 7-point scale ranges from 1 = "completely inapplicable" to 7 = "completely applicable." The measure meets all conventional reliability standards with Cronbach's alpha ranging from 0.77 to 0.86 (see Section A.1 in Online Appendix A, available as supplemental material).

⁷Because several workers refused to do the vignette task, whereas others provided the maximum of 48 responses, Tobit regressions constitute an alternative estimation method in such cases with censored data. Using this estimation technique does not qualitatively alter the results.

⁸Adding the number of days between completing the previous job and receiving updated information to either specification, however, there is no evidence of such a negative relationship—the estimated coefficients are insignificant and even positive in sign. Allowing for a nonlinear relationship between the time passed and the number of answered vignettes by including dummy variables for the number of weeks already passed has no effect either.

⁹We also modified this measure of vignette responses by applying a quality adjustment to eliminate unsuitable answers ("do not know" and "not applicable" answers). Using this approach does not alter results qualitatively.

¹⁰Score on a 5-point scale, ranging from 1 = "very slightly or not at all" to 5 = "very much."

¹¹For statistics on all emotional items, see Section A.7 in Online Appendix A.

¹²Estimating the impact of treatments on participants' emotional state using regression analysis (ordered probit models) with the full set of controls (sociodemographics, Big Five, and organizational controls) does not alter the results (*p*-values are ranging from 0.002 to 0.075).

References

- Akerlof GA (1982) Labor contracts as partial gift exchange. *Quart. J. Econom.* 97(4):543–569.
- Akerlof GA, Yellen JL (1990) The fair wage-effort hypothesis and unemployment. *Quart. J. Econom.* 105(2):255–283.
- Al-Ubaydli O, List JA (2015) On the generalizability of experimental results in economics. Fréchet GR, Schotter A, eds. *Handbook of Experimental Economic Methodology* (Oxford University Press, New York), 420–462.
- Ariely D, Kamenica E, Prelec D (2008) Man's search for meaning: The case of Legos. *J. Econom. Behav. Organ.* 67(3–4):671–677.
- Ashraf N, Bandiera O, Jack BK (2014) No margin, no mission? A field experiment on incentives for pro-social tasks. *J. Public Econom.* 120:1–17.
- Bäker A, Mechtel M (2013) The role of task meaning on output in groups: Experimental evidence. IAAEU Discussion Paper 08/2013, Institute for Labour Law and Industrial Relations in the European Union, Trier University, Trier, Germany.
- Becker SO, Messer D, Wolter SC (2013) A gift is not always a gift: Heterogeneity and long-term effects in a gift exchange experiment. *Economica* 80(318):345–371.
- Berg J, Dickhaut J, McCabe K (1995) Trust, reciprocity, and social history. *Games Econom. Behav.* 10(1):122–142.
- Bradler C, Dur R, Neckermann S, Non A (2016) Employee recognition and performance: A field experiment. *Management Sci.* 62(11):3085–3099.
- Carpenter J, Gong E (2016) Motivating agents: How much does the mission matter? *J. Labor Econom.* 34(1):211–236.
- Chandler D, Kapelner A (2013) Breaking monotony with meaning: Motivation in crowdsourcing markets. *J. Econom. Behav. Organ.* 90:123–133.
- Compte O, Postlewaite A (2004) Confidence-enhanced performance. *Amer. Econom. Rev.* 94(5):1536–1557.
- Costa PT, McCrae RR (1989) The NEO PI/FFI Manual supplement (Psychological Assessment Resources, Odessa, FL).
- Falk A, Heckman JJ (2009) Lab experiments are a major source of knowledge in the social sciences. *Science* 326(5952):535–538.
- Falk A, Ichino A (2006) Clean evidence on peer effects. *J. Labor Econom.* 24(1):39–57.
- Fehr E, Kirchsteiger G, Riedl A (1993) Does fairness prevent market clearing? An experimental investigation. *Quart. J. Econom.* 108(2):437–459.
- Fehr E, Kirchsteiger G, Riedl A (1998) Gift exchange and reciprocity in competitive experimental markets. *Eur. Econom. Rev.* 42(1):1–34.
- Fehrler S, Kosfeld M (2014) Pro-social missions and worker motivation: An experimental study. *J. Econom. Behav. Organ.* 100:99–110.
- Fried Y, Ferris GR (1987) The validity of the job characteristics model: A review and meta-analysis. *Personnel Psych.* 40(2):287–322.
- Gächter S, Fehr E (2002) Fairness in the labour market. Bolle F, Lehmann-Waffenschmidt M, eds. *Surveys in Experimental Economics Bargaining, Cooperation and Election Stock Markets* (Physica-Verlag, Heidelberg, Germany), 95–132.
- Gneezy U, List JA (2006) Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica* 74(5):1365–1384.
- Grant AM (2008) The significance of task significance: Job performance effects, relational mechanisms, and boundary conditions. *J. Appl. Psych.* 93(1):108–124.
- Hackman JR, Oldham GR (1976) Motivation through the design of work: Test of a theory. *Organ. Behav. Human Performance* 16: 250–279.
- Harrison GW, List JA (2004) Field experiments. *J. Econom. Literature* 42(4):1009–1055.
- Hennig-Schmidt H, Sadrieh A, Rockenbach B (2010) In search of workers' real effort reciprocity—A field and a laboratory experiment. *J. Eur. Econom. Assoc.* 8(4):817–837.
- Humphrey SE, Nahrgang JD, Morgeson FP (2007) Integrating motivational, social, and contextual work design features: A meta-analytic summary and theoretical extension of the work design literature. *J. Appl. Psych.* 92(5):1332–1356.
- Jamison J, Karlan D, Schechter L (2008) To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments. *J. Econom. Behav. Organ.* 68(3–4):477–488.
- Jeworrek S, Mertins V (2014) Wage delegation in the field. IAAEU Discussion Paper 08/2014, Institute for Labour Law and Industrial Relations in the European Union, Trier University, Trier, Germany.
- Kosfeld M, Neckermann S (2011) Getting more work for nothing? Symbolic awards and worker performance. *Amer. Econom. J.: Microeconomics* 3(3):86–99.
- Kosfeld M, Neckermann S, Yang X (2017) The effects of financial and recognition incentives across work contexts: The role of meaning. *Econom. Inquiry* 55(1):237–247.
- Krohne HW, Egloff B, Kohlmann CW, Tausch A (1996) Untersuchung mit einer deutschen Version der "Positive and Negative Affect Schedule" (PANAS). *Diagnostica* 42(2):139–156.

- Kube S, Maréchal MA, Puppe C (2012) The currency of reciprocity—Gift-exchange in the workplace. *Amer. Econom. Rev.* 102(4):1644–1662.
- Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? *J. Econom. Perspect.* 21(2):153–174.
- Loewenstein G (1999) Because it is there: The challenge of mountaineering for utility theory. *Kyklos* 52(3):315–343.
- Maniadis Z, Tufano F, List JA (2014) One swallow doesn't make a summer: New evidence on anchoring effects. *Amer. Econom. Rev.* 104(1):277–290.
- Rick S, Loewenstein G (2008) The role of emotion in economic behavior. Lewis M, Haviand Jones J, Feldmann Barett L, eds. *Handbook of Emotions*, 3rd ed. (Guilford, New York), 138–156.
- Rosso BD, Dekas KH, Wrzesniewski A (2010) On the meaning of work: A theoretical integration and review. *Res. Organ. Behav.* 30:91–127.
- Searle R, Den Hartog DN, Weibel A, Gillespie N, Six F, Hatzakis T, Skinner D (2011) Trust in the employer: The role of high-involvement work practices and procedural justice in European organizations. *Internat. J. Human Resource Management* 22(5):1069–1092.
- Tonin M, Vlassopoulos M (2010) Disentangling the sources of pro-socially motivated effort: A field experiment. *J. Public Econom.* 94(11–12):1086–1092.
- Watson D, Clark LA, Tellegen A (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Personality Soc. Psych.* 54(6):1063–1070.
- Weller I, Matiaske W (2009) Persönlichkeit und Personalforschung. Vorstellung einer Kurzskaala zur Messung der "Big Five." *Zeitschrift für Personalforschung* 23(3):258–266.
- Zizzo DJ (2010) Experimenter demand effects in economic experiments. *Experiment. Econom.* 13(1):75–98.