

Can lab experiments help design personnel policies?

Employers can use laboratory experiments to structure payment policies and incentive schemes

Keywords: laboratory experiments, behavior, incentives, evaluation, peer effects, human resource management

ELEVATOR PITCH

Can a company attract a different type of employee by changing its compensation scheme? Is it sufficient to pay more to increase employees' motivation? Should a firm provide evaluation feedback to employees based on their absolute or their relative performance? Laboratory experiments can help address these questions by identifying the causal impact of variations in personnel policy on employees' productivity and mobility. Although they are collected in an artificial environment, the qualitative external validity of findings from the lab is now well recognized.

KEY FINDINGS

Pros

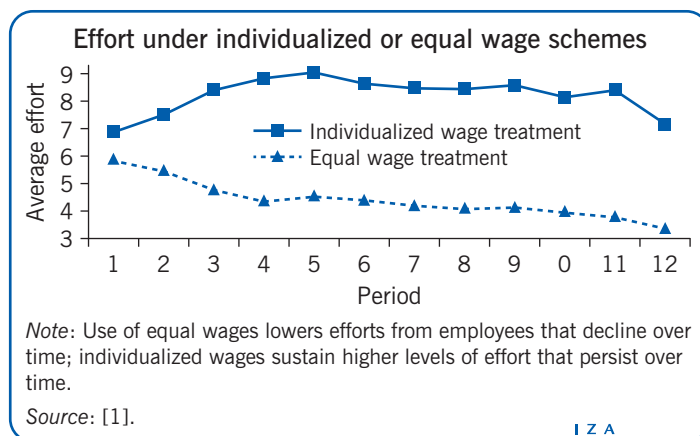
- + The laboratory offers a cheap and reliable method to identify causal relationships between incentives and effort, with a high qualitative external validity.
- + The anticipated impact of a change in compensation policies on employees' effort and sorting can be identified in the laboratory.
- + Lab experiments can determine how reference points in terms of fair wage, comparison income, expectations, and goals influence the intensity of effort.
- + The importance of simultaneously studying the monetary and non-monetary sources of motivation to better understand how they interact can be demonstrated in a laboratory context.
- + The laboratory provides clean measures of peer effects on productivity and information on optimal performance feedback policies.

Cons

- Subjects are observed only during a brief period of time; it is thus difficult to design lab experiments that are able to capture the long-term effects of a change.
- Subjects participating in lab experiments self-select and may feel scrutinized, with consequences on their behavior.
- Students and CEOs very probably have a number of different and important characteristics, but it is difficult to involve CEOs in lab experiments.

AUTHOR'S MAIN MESSAGE

Many companies try to improve their human resources management as a way to increase productivity and offer a working environment that is able to attract and retain the best employees. Estimating the potential impact of a new incentive scheme is, however, very difficult. Laboratory experiments can help anticipate the costs and benefits of a change by investigating the economic, social, and psychological aspects of human reactions to positive and negative incentives. While there are legitimate questions about the extent to which laboratory findings are applicable to the wider population, recent evidence is reassuring.



MOTIVATION

Companies always need to identify new sources of productivity, and one way of doing so is to design appropriate incentives, as this is crucial to motivate employees. Further, incentives that are able to strike a balance between competition and cooperation between employees can certainly offer a comparative advantage to the firm. However, this requires an understanding by the firm of the complexity of human motivation. For example, offering a higher wage than competitors may not be sufficient to attract the desired workforce if people also care about wage inequality and non-wage sources of motivation. In addition, the way in which people communicate job-related information has changed enormously in recent years: social networks are now a principal channel for how people evaluate their job satisfaction and set their career goals. This results in a higher degree of scrutiny of, and responsiveness to, different sources of motivation.

Insights from behavioral economics on social preferences (e.g. reciprocity, altruism, inequality aversion) and cognitive biases (e.g. over- and under-confidence, biased updating of beliefs, preference for the present) are fundamental to improve a firm's understanding of how workers may react to incentives. Standard surveys do not usually incorporate such measures. However, because laboratory experiments in labor economics allow researchers to collect these data and control incentives, they can provide a powerful instrument to help explain why existing policies are sometimes effective and sometimes not, particularly when efficiency is affected by imperfect information and strategic interactions. Recent research provides reassuring evidence regarding the validity of laboratory experiment results “in the real world.”

DISCUSSION OF PROS AND CONS

A lab experiment generates data in a controlled environment by means of a protocol that reproduces the properties of a theory, with subjects (usually students) making anonymous decisions in the presence of monetary incentives, without using deception. Varying treatments allows researchers to identify, at low cost and in a controlled environment, the impact of external changes in the institutions regulating the exchanges between players. While non-experimental labor economics research mainly focuses on the adjustment processes in the labor market, lab experiments have shed light on how and when incentives work. Major insights are illustrated below, in four domains.

The incentive and sorting effects of compensation schemes

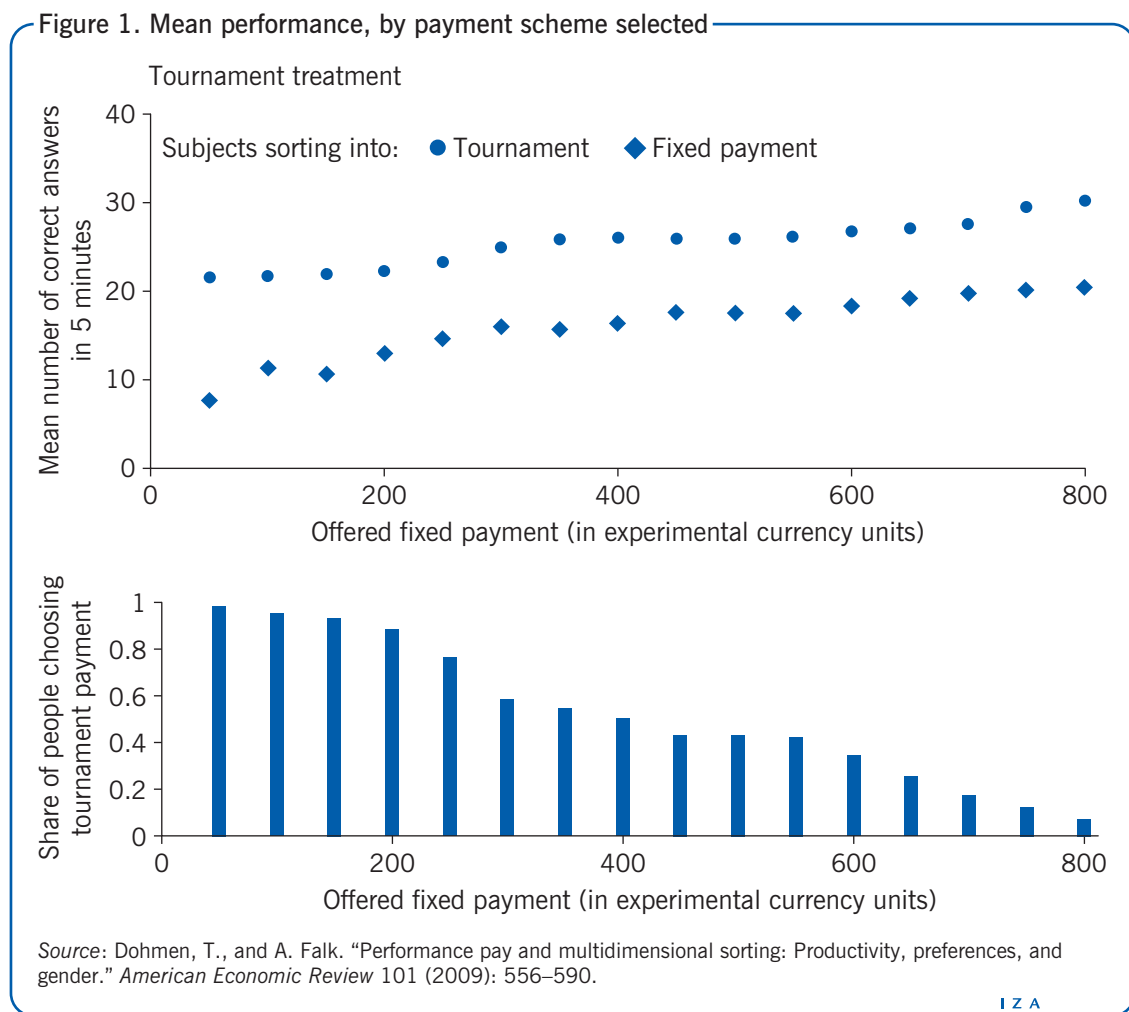
Different compensation schemes give rise to different levels of effort from employees. Contrary to theoretical predictions based on the “selfish” preferences of workers, when labor contracts are incomplete (in the sense that these contracts cannot formally impose a certain level of productivity corresponding to each state of the world), a higher wage can encourage employees to provide more than their minimum effort if they value the potential benefits of their higher effort for the firm that offered a higher wage. Indeed, this sort of reciprocal response from the employee serves as a contract enforcement device, that is, it motivates employees to exert effort even when contracts are incomplete [2]. Reciprocity, in this sense, supports a “gift exchange” relationship between employers and employees, in that employees will reciprocate a generous wage offer by providing a high level of effort.

Over time, even the more selfish employees will mimic the reciprocal ones, supporting voluntary cooperation and the emergence of long-term employment relationships.

Laboratory experiments have also shown that revenue-sharing between team members, regardless of their individual level of effort, does not necessarily induce free-riding among them, especially when team members are able to self-govern as a result of communication and peer pressure.

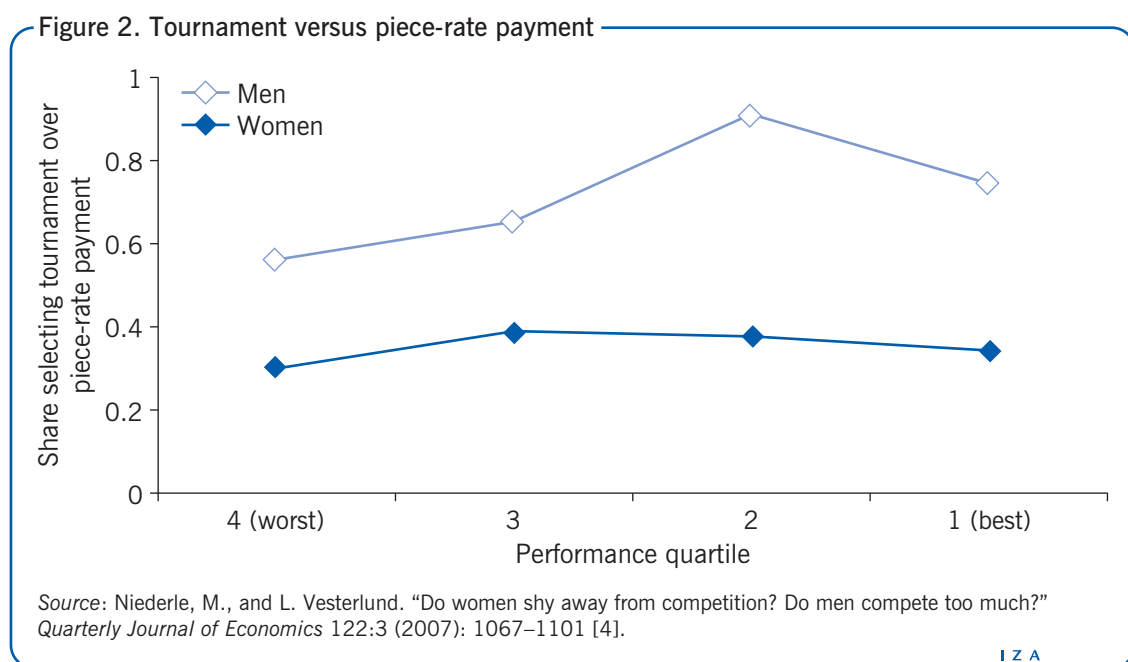
The introduction of performance pay, such as piece-rates, can increase performance compared to fixed wages, as individuals react strongly to monetary incentives (Figure 1). Competitive payment schemes, such as tournaments, can increase performance even further, except when the winner of the competition can earn only slightly more than the loser. This comes, however, at the price of a higher variance in the levels of effort and with risks of sabotage and ingratiating or rebellious behavior that can reverse the motivating effect of competition.

In the real world, simply correlating a new compensation scheme with average productivity, without taking into consideration the fact that employees are mobile between firms and can thus choose their payment scheme to some extent, would deliver misleading conclusions on the incentive effect of this scheme. Indeed, changes in compensation rules can have not only incentive effects but also sorting effects [3]. For example, the introduction of performance pay will be attractive for individuals who feel able to improve



their performance, but repulsive for workers who cannot, or are not willing to increase their effort. The pure incentive effect of a compensation scheme can be identified only once its impact on the hiring and exit of employees has been measured. Mobility is welfare-improving because it allows people to select their payment scheme, which increases their “utility.” For example, risk-averse individuals tend to dislike payment schemes based on competition between employees; when these individuals are forced to compete, they usually exert an excessively costly level of effort that may nevertheless not allow them to succeed. If they can choose, they will self-select in a different payment scheme, that is, they will search for a job in firms offering other payment schemes. As a result, the variance of effort in companies using tournaments diminishes, which increases efficiency.

The various incentive schemes are perceived differently according to gender, age, risk attitudes, cognitive abilities, personality traits, and preferences [2]. Part of the explanation of the gender gap in access to top corporate positions, for instance, relies on females’ lower willingness to opt into competitive compensation schemes. Even the most able women tend to shy away from competition (Figure 2) [4]. This gender gap in competitiveness remains even after controlling for differences in risk attitudes, self-confidence, and the aversion to receiving feedback on one’s performance. Introducing quotas for access to top positions would thus improve efficiency, as it convinces the most able women to enter and win the competition without passing over the most competent men. In contrast, women are more attracted than men by revenue-sharing in teams.



The importance of reference points

Employees’ motivation depends on “reference points,” either in terms of wages or in terms of performance. The notion of a reference point is associated with that of loss-aversion: individuals suffer more “disutility” below this reference point than the utility they gain when they exceed this point by a similar amount. An important reference point in the determination of effort is what people consider to be a fair wage. This fair wage may be defined both in absolute and in relative terms. In particular, from the perspective of

equity, merit pay is important when skills differ. While employers tend to compress the wage distribution (i.e. reduce salary differentials), especially when there is uncertainty about the employees' skills, higher-ability employees tend to reduce effort when they are not paid more than their co-workers of lower ability. The same detrimental effect is found when firms are forced to pay the same wage to employees providing different levels of effort, as is frequently the case for example in the public sector.

When skills are similar amongst employees, effort is affected by discriminatory relative wages. This means that for a given wage, employees will provide a lower effort when the average income of their reference group is higher, and a higher effort if their reference group members earn less on average [5]. This is also observed when individuals compare themselves to a position they may have held in the past that paid more: demotions (in terms of relative position) have a stronger impact on current effort than promotions. This effect can be referred to as “loss aversion,” whereby individuals suffer more from getting an outcome below the reference point than they feel happy when their outcome exceeds this point. Similarly, in long-term relationships, when contracts have to be renegotiated, the initial contract usually constitutes a strong reference point.

The reference point may also be the individuals' expectations about their future [6]. Expectations as a reference point and income targeting can thereby explain apparent anomalies. For example, sometimes people do not work more even when wages increase. Instead, individuals may substitute leisure for work once they have obtained a certain target income, even though they could earn much more by working longer hours. Expectations and income targeting may also explain why some unemployed people search too long for a job when they have unrealistic expectations about their ability to secure a job that offers this target income, given the economic environment.

Goal-setting, especially when it is not linked to salary increases, introduces another type of reference point in the determination of effort. Letting employees set their own productive goals increases the intrinsic motivation of those who have a sense of self-achievement, which supports adapting contracts to the type of employee. The effect is stronger than when goals are set externally. The “goal gradient theory,” developed by the psychologist Clark L. Hull, explains that effort intensifies when the goal gets closer. Experimental tests of this theory have shown that individuals adjust their effort according to the distance between their current performance and their goal. Some workers may use goal-setting to commit themselves when they know they suffer from self-control problems and a tendency to stop effort too quickly.

Non-monetary sources of motivation

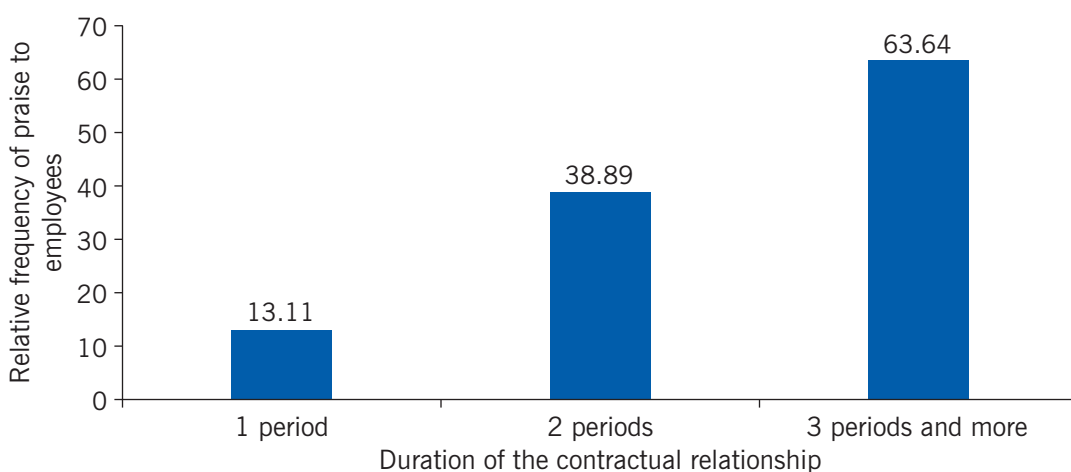
While psychologists traditionally focus on intrinsic motivation (personal achievement) and economists on extrinsic motivation (incentives), behavioral economics and laboratory experiments have shown how monetary and non-monetary incentives interact. While increasing material incentives usually improves motivation, many lab experiments have revealed that the manipulation of these incentives is very complex. This is because changing the incentives may change the perceptions of employees of the nature and meaning of the task. In some situations, the introduction of monetary incentives “crowds out” the intrinsic motivation to perform a task [7]. Performance may decrease after the introduction of incentives, for several possible reasons: (i) the incentives harm self-image by giving the impression that the individual is willing to perform only for money; (ii) it

sends a signal that the task is particularly difficult to perform; or (iii) the nature of the employment relationship is modified (i.e. the service becomes a commodity). In other situations, increasing the level of incentives by a large proportion generates excessive motivation that may have detrimental consequences on performance (“choking under pressure”).

Similar results have been found on disincentives, that is, the introduction of monitoring and associated sanctions. For example, reducing the workers’ set of possible actions by not allowing them to choose any level of effort has a depressing effect on their performance. Thus, although one cannot exclude a disciplining effect of monitoring under many circumstances, an excessively high level of monitoring discourages effort when the employment relationship is repeated and interpersonal. Instead, individuals seem to more readily accept pressure coming from peers than from the hierarchy. Meanwhile, individuals try to retain authority even when payoffs could be increased by delegating more, which leads them to over-provide effort while subordinates tend to under-provide effort [8].

While material incentives can sometimes harm performance, non-monetary incentives, such as status, can support effort. The assignment of symbolic rewards (e.g. “employee of the month,” “thank you” e-cards) or symbolic sanctions from peers, encourages workers to perform better, even though they are “payoff-irrelevant” [9]. Employee engagement is also increased when the meaning of his or her task is made clear, and by the expression of employers’ recognition in private. However, to be effective in the employees’ eyes, praise must be costly in some way to the employer, indicating that the employer is worth impressing: in other words, praise is valuable only when it comes from people one respects. Moreover, compared to short-term contracts, longer relational contracts are more likely to have been initiated by the assignment of symbolic rewards praising the employee for his or her performance (Figure 3) [10].

Figure 3. Relational contracts and symbolic rewards



Note: Shows the percentage of employer-employee relationships initiated by symbolic rewards, by total duration (number of experimental periods) of the relationship. At the end of each experimental period, employers can assign symbolic rewards to praise employees for their effort, at a cost to themselves. At the beginning of each new period, the labor market reopens and firms have to make wage offers to hire workers. The figure shows that the assignment of symbolic rewards facilitates the reemployment of the same workers over time because it expresses trust on both sides. For example, in only 13.11% of the contracts that end after one period of interaction have employers expressed praise to their employee; in contrast, this is the case in 63.64% of the contracts that are renewed at least three times.

Source: Eriksson, T., and M. C. Villeval. “Respect and relational contracts.” *Journal of Economic Behavior & Organization* 81 (2012): 286–298 [10].

The importance of non-monetary rewards and respect comes from the fact that people care not only about what others *do* to them but also what others *think* about them. This can be rationalized by defining the performance motive of individuals as including monetary incentives, intrinsic motivation, as well as reputational concerns when actions are visible to others.

Furthermore, workers rarely perform in isolation. If it is difficult for them to compare their wages to that of their co-workers because of pay secrecy, they are usually able to compare their productivity to that of their peers because they can observe each other, or because they receive feedback. Indeed, a question for companies is whether or not to provide employees with feedback and whether feedback should be on absolute or on relative performance. This also has significant implications for the design of work environments.

Most laboratory experiments have found that providing subjects with information about their performance relative to that of other relevant individuals, or letting people observe co-workers, increases their motivation. This is especially true for the less-able workers, who benefit from the presence of peers [11]. These positive peer effects on productivity have been found in settings in which compensation schemes are either flat, or based on piece-rates or on tournaments. Investigations of the mechanisms behind these peer effects show the role of conformity, competitiveness, disutility from lagging behind, and peer pressure (when being observed by more-able workers). The presence of peers also has an influence on quitting behavior: peers are more likely to work longer with peers present than when working alone, and tend to quit at a similar time when they have the right to communicate, mainly because of a higher sociability of the environment.

It is important to note, however, that sometimes no effect, or a negative effect, of peers' performance on individual effort has been found. The adverse impact is more pronounced for lowest-performers and disappointment-averse competitors in tournaments. Work quality may also be harmed by peer comparisons: indeed, peer effects may increase effort provision but also the rate of defects.

Finally, a gender effect has been found when peers interact within networks. Both women and men increase their performance in line with their peers when those peers are not informed of their performance, as if peers help set the goal. In contrast, women disregard information on peers when they interact simultaneously with these peers instead of when they receive information about the past performance of peers, as if the situation was perceived as being more competitive.

It should be noted that laboratory experiments provide very robust tools to identify peer effects. In the business setting, when a worker and his co-workers' behaviors are similar, survey data usually do not permit the identification of whether this is due to workers having similar characteristics, or being affected by similar shocks, or influenced by peers' behavior. The presence of payoff interdependencies in the compensation scheme also makes the identification of peer effects more difficult outside of the lab. Importantly, comparisons between laboratory and field experiments indicate that the estimated coefficients of peer effects do not differ significantly between the laboratory and the field (the "real world"). Findings from the lab generalize fairly well to the field, both qualitatively and quantitatively.

The perceived “generalizability” of laboratory findings

The external validity of laboratory experiments—that is, the possibility of extrapolating findings on causality, etc., from the lab to the real world—is a topic of buoyant discussion. A number of threats to generalizability have been identified. Behavior in the lab may deviate from the course of action that the same individuals would follow in their real life because they may feel “under scrutiny” in the lab. A lack of perfect anonymity between subjects and the experimenter may also bias behavior. Moreover, the context in which a decision is embedded may be very specific, which is not entirely controlled by the experimenter, and small variations in the instructions can make big differences in behavior. In addition, subjects who volunteer in experiments may have specific characteristics: for example, they are drawn from a specific group, such as students, who are not necessarily representative of the whole population. Furthermore, most laboratory experiments are very brief, whereas many employment relationships usually involve long-term interactions. Finally, the monetary stakes of the decisions in the lab are small and this may challenge the knowledge on whether similar behavior would hold when confronted by much bigger stakes.

Reassuringly, however, studies comparing the behavior of student subjects with the behavior of professionals, workers, or CEOs, placed in the same experimental conditions, typically do not find differences, except that students tend to behave more selfishly. There are a number of similar results for both types of subjects: (i) an increase in the number of competitors in a laboratory tournament reduces effort; (ii) in repeated interactions, individuals try to conceal their true ability when they anticipate that revealing it will lead to an increase in the standards imposed on workers in the future; (iii) effort for a given wage increases with the rank in the wage distribution. This does not mean that students and CEOs have the same characteristics (CEOs may be more competitive and less risk-averse), but they react in a similar way to the manipulation of treatments in the laboratory compared to what theory predicts.

Moreover, most criticism directed at the method can be addressed by the experimental method itself. For example, raising the level of monetary stakes in lab experiments does not change the degree of players’ pro-sociality: people still show reciprocity, trust, or altruism even when their costs increase significantly.

Notwithstanding the fact that the criticism on generalizability can also apply to field experiments, econometric case studies, and even survey data, the advantages of laboratory methods are substantial. The replicability of the experiments, the possibility to introduce exogenous changes, and the control conditions, allow researchers to test the impact of modifications in the environment. Overall, there is a consensus in the research community that laboratory experiments have a qualitative external validity and can deliver reasonable, directional, real-world predictions.

Whether they can make precise point predictions about the magnitude of an effect remains a topic for discussion. However, this criticism can be applied to many empirical studies. Moreover, some recent evidence has shown that the size of peer effects on individual productivity is not significantly different when estimated in the laboratory and in the field.

LIMITATIONS AND GAPS

Several important dimensions remain to be explored. Compared to the role of social preferences in the determination of wage policies and effort decisions, behavioral labor economics has investigated the importance of cognitive biases to a much lesser extent. Although there are a few studies on over-confidence and procrastination, developing this approach could help understand how people distort their beliefs about their ability and their chance of success, which is an important source of motivation. Lab experiments, including neuro-economic studies, could provide major insights on the formation of motivated beliefs, with implications in search behavior on the labor market and employment relationships within the firm.

As far as biases are concerned, lab experiments have focused on individual employees. However, little is known on the biases existing on the firms' side and on how they develop in groups. When does group decision-making help eliminate biases—if “two heads are better than one”—and when does it amplify individual mistakes—when motivated beliefs lead to ignoring relevant signals, which in turn potentially leads to catastrophes? And how does group decision-making relate to group identity?

Further experiments are needed to better understand how new modes of organizations, like networks, can generate novel forms of cooperation, peer pressure, or leisure-work arbitrage; but also to better understand how to incentivize the dimensions of collective intelligence required by innovation, like creativity. When peer effects are studied, the existence of a hierarchy within the network is usually ignored. It would be important to test whether individuals react more to the presence of managers or to the influence of peers, conditional on the type of organization [12], and how the style of management (monitoring vs coaching) impacts motivation. This inclines in favor of developing experiments that study the efficiency of mechanisms supporting cooperation at the intersection between labor economics and organizational behavior.

Finally, one important limitation of current lab experiments is that they are not very well equipped to take into account long-term horizons. This is also why lab experiments, field experiments, and survey data are needed as complements and should certainly not be seen as competing methods.

SUMMARY AND POLICY ADVICE

Laboratory experiments have provided major insights on when and why incentives work or not. If individuals are, in general, motivated by the prospect of higher monetary gains, many other dimensions matter. Social preferences, such as reciprocity or inequality aversion, influence the level of involvement, the nature and the duration of contracts, and the self-selection of individuals into different payment schemes.

Social comparisons and moral norms, in terms of justice and equity, also affect the judgments of individuals on how fairly they are treated compared to co-workers. Envy, competitiveness, or mimicry also influence how hard individuals are willing to work to minimize the distance between their performance relative to peers. Risk attitudes, expectations, and cognitive processes matter when people decide on how much effort to invest, especially when the environment and incentives are uncertain. Reference points influence the goals individuals set for themselves. Overall, these factors explain how complex the manipulation of incentives remains.

On the downside, to produce reliable findings, laboratory experiments must simplify as much as possible the work environment in order to be able to identify causal relationships between incentives and effort. In contrast, the real world of labor is much more complex, notably due to dynamics and to its embeddedness in economic and institutional frameworks. Such a complexity cannot (and should not) be introduced in the lab. This leads to two conclusions. First, it should be kept in mind that the lab does not intend to mimic the real world but the conditions of theoretical models, with the purpose of testing their behavioral predictions. Second, the insights from lab experiments need to be complemented by other sources of empirical investigations, such as field experiments and survey analyses.

Policymakers have to make decisions in very uncertain environments. Reducing this uncertainty by trying to identify regularities is helpful. The laboratory should be considered as not only an academic instrument, but also as a test bed to design renewed personnel policies.

Acknowledgments

The author thanks an anonymous referee and the IZA World of Labor editors for many helpful suggestions on earlier drafts. Previous work of the author contains a larger number of background references for the material presented here and has been used intensively in all major parts of this article [5], [10]. This work was performed within the framework of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

Competing interests

The IZA World of Labor project is committed to the *IZA Guiding Principles of Research Integrity*. The author declares to have observed these principles.

© Marie Claire Villeval

REFERENCES

Further reading

- Charness, G., and P. J. Kuhn. "Lab labor: What can labor economists learn from the lab?" In: Ashenfelter, O., and D. Card (eds). *Handbook of Labor Economics, Volume 4*. Amsterdam: Elsevier, 2011; pp. 229–330
- Dohmen, T. "Behavioral labor economics: Advances and future directions." *Labour Economics* 30 (2014): 71–85.
- Fréchette, G. R. "Laboratory experiments: Professionals versus students." In: Fréchette, G. R., and A. Schotter (eds). *Handbook of Experimental Economic Methodology*. Oxford: Oxford University Press, 2015; pp. 360–390.

Key references

- [1] Abeler, J., S. Altmann, S. Kube, and M. Wibral. "Gift exchange and workers' fairness concerns: When equality is unfair." *Journal of the European Economic Association* 8:6 (2010): 1299–1324.
- [2] Fehr, E., S. Gächter, and G. Kirchsteiger. "Reciprocity as a contract enforcement device: Experimental evidence." *Econometrica* 65:4 (1997): 833–860.
- [3] Lazear, E. P. "Performance pay and productivity." *American Economic Review* 90:5 (2000): 1346–1361.
- [4] Niederle, M., and L. Vesterlund. "Do women shy away from competition? Do men compete too much?" *Quarterly Journal of Economics* 122:3 (2007): 1067–1101.
- [5] Clark, A., D. Masclet, and M. C. Villeval. "Effort and comparison income: Experimental and survey evidence." *Industrial and Labor Relations Review* 63:3 (2010): 407–426.
- [6] Abeler, J., A. Falk, L. Goette, and D. Huffman. "Reference points and effort provision." *American Economic Review* 101:2 (2011): 470–492.
- [7] Gneezy, U., S. Meier, and P. Rey-Biel. "When and why incentives (don't) work to modify behavior." *Journal of Economic Perspectives* 25:4 (2011): 191–209.
- [8] Fehr, E., H. Herz, and T. Wilkening. "The lure of authority: Motivation and incentive effects of power." *American Economic Review* 103:4 (2013): 1325–1359.
- [9] Kosfeld, M., and S. Neckermann. "Getting more work for nothing? Symbolic awards and worker performance." *American Economic Journal: Microeconomics* 3:3 (2011): 86–99.
- [10] Eriksson, T., and M. C. Villeval. "Respect and relational contracts." *Journal of Economic Behavior and Organization* 81:1 (2012): 286–298.
- [11] Falk, A., and A. Ichino. "Clean evidence on peer effects." *Journal of Labor Economics* 24:1 (2006): 39–57.
- [12] Lazear, E. P., K. Shaw, and C. Stanton. *The Value of Bosses*. NBER Working Paper No. 18317, 2012.

Online extras

The **full reference list** for this article is available from:

<http://wol.iza.org/articles/can-laboratory-experiments-help-design-personnel-policies>

View the **evidence map** for this article:

<http://wol.iza.org/articles/can-laboratory-experiments-help-design-personnel-policies/map>