

# Experimental economics

## Lecture 6: Interactions in the experiment

Matej Lorko

matej.lorko@euba.sk

Materials: [www.lorko.sk/lectures](http://www.lorko.sk/lectures)

### References:

- Weimann, J., & Brosig-Koch, J. (2019). *Methods in experimental economics*. Springer International Publishing. Chicago
- Jacquemet, N., & l'Haridon, O. (2018). *Experimental economics*. Cambridge University Press.

# The Experimenter Demand Effect

- The experimenter influences what happens in an experiment through different channels. Some are obvious, such as the instructions given to the subjects by the experimenter, or the exercises used to test whether the subjects have understood the experiment. Others are less obvious, but just as important. Thus, the experimenter can consciously or unconsciously exert social pressure or certain expectations can be generated in the subjects as to the purpose of the experiment and what behavior is now expected of them.
- In laboratory experiments the interaction between the experimenter and the subject is inevitable (even if it is through the design developed by the experimenter). It cannot therefore be a question of avoiding any kind of interaction, but rather of designing it in such a way that it does not lead to any distorting influence on the behavior of the subjects (experimenter demand effect), thereby curtailing the interpretability of the data obtained.
- Cognitive experimenter demand effects occur because the experimenter has to explain the experiment to the subjects. Understanding this explanation is a cognitive process and it may well happen that how it is explained leads to it being understood in a particular way, for example what is appropriate behavior in the experimental situation. Experimenters should be aware of the fact that subjects may take every word seriously and, therefore, that every word used by the experimenter should be carefully considered.
- In addition to cognitive experimenter demand effects, undesirable manipulation of the subjects may also result from social pressure, which can arise both between the subjects and vertically from the experimenter. There are many reasons why people succumb to social pressure. For example, a role may be played by the desire for conformity, or by social acceptance, which is experienced when acting in accordance with a social norm. It is quite possible that there are also subjects who attach great importance to being nonconformist and therefore oppose any social pressure. While it may not be too bold a hypothesis to suggest that nonconformists are rare, the widespread desire to conform to social norms is well known.

# The Experimenter Demand Effect

- The instructions that the subjects receive at the beginning of an experiment are ideally suited to creating massive experimenter demand effects. The language used, for example, is suspected of doing this. It is possible to describe things in an emphatically neutral way or to “load” them with valuations to a greater or lesser extent.
- Liberman et al. (2004) report on two public good experiments, which were identical except for the names of the games provided to the subjects. One was a “Community Game” and the other was a “Wall Street Game”. The names actually had a huge influence on the results, with much more cooperation in the Community Game than in the Wall Street Game.
- In the experiment by Burnham et al. (2000), too, altering only one word triggered substantial effects. In their experiment, two players could significantly increase their payoffs compared to the equilibrium payoff if player 1 trusted player 2 and player 2 acted reciprocally, thus vindicating the trust. In the first treatment, the other player was called the “partner”, while in the second treatment the word “opponent” was used. The word “partner” led to significantly more trust and trustworthiness at the beginning of the experiment. Admittedly, both declined in later rounds.
- The decisive question in both cases is what effect is actually present. Is it a particular value judgment associated with the respective terms, or is it an experimenter demand effect? In the latter case, when a game is called “Wall Street Game”, the subjects might have the feeling that the experimenter wants to test how well they can assert themselves. If the game is called “Community Game”, the experimenter might want to know how well the subjects perform as social beings. If the other player is called a partner, the experimenter apparently wants to test the ability to cooperate. If, on the other hand, the other player is designated an “opponent”, then competition is evidently at issue and it is a matter of asserting oneself.
- Experimenter demand effects can act in different directions. The reference point is the experimental effect expected in the experiment. The experimenter demand effect may be in the same direction, opposite or orthogonal to the experimental effect.
- The most problematic is the experimenter demand effect that acts in the same direction as the expected experimental effect. In such a case, it is difficult to decide whether what is observed is due to the experimenter demand effect or to the experimental conditions. If the experimenter demand effect runs in the opposite direction, it can just offset the experimental effect and no clear effects can be detected. The least problematic are experimenter demand effects that are orthogonal to the experiment effect. They may not influence the behavior of the subjects in a way that hinders the interpretation of the results of the experiment.

# Double-Blind Design

- Double-blind procedure is an experimental design that ensures that the experimenters cannot observe how the individual subject acts and that also maintains anonymity between the subjects. This is generally achieved by having the subjects drawing identification numbers randomly and in a concealed manner. As a result, the experimenters know how, for example, subject number 17 behaved, but not who number 17 is. A single-blind design means that the subjects cannot observe each other, but the experimenter sees what the individual person is doing.
- It is essential to see double-blind designs in close conjunction with the experimenter demand effect. This is necessary because it cannot be ruled out that the use of a double-blind design itself will trigger an experimenter demand effect. If experimenters explicitly draw the attention of their subjects to the fact that they are acting anonymously and cannot be observed by the experimenter, then it is obvious that the subjects will think about why it is so important to the experimenter that they can act without being observed. Therefore, when using a double-blind design, it is not advisable to explicitly point out that this is intended to achieve anonymity.
- Double-blind designs are particularly effective where a strong experimenter demand effect is expected, however, if a sufficiently high degree of anonymity is already guaranteed by a single-blind design, double-blind might not be necessary.

# The Frame of the Experiment

- The frame of an experiment is the way in which a specific decision problem is presented to the subjects. Framing effects are the changes in the subjects' behavior that occur solely because the presentation of the decision problem is varied without changing the problem itself and its solution.
- In the recent literature, two types of framing effects play a special role. The first occurs when only the name of a game is changed (label frame). We have already referred to the following example in the previous section. Whether you call a public good experiment “Community Game” or “Wall Street Game” makes a major difference.
- The second framing effect that has attracted much attention is what is named the valence frame. This means that certain terms are loaded with respect to the values or preconceptions associated with them. The standard example again concerns the public good game, which can be played in a “Give” or a “Take” treatment (Dufwenberg et al. 2011).
- In the Give frame, the individual members of a group each receive an initial endowment ( $z_i$ ), which they can either keep or pass on to any part of a joint project (the public good). In the Take frame, the entire initial endowment (i.e. the sum of the  $z_i$ ) is in the joint project and the subjects can withdraw money up to the amount of  $z_i$ . Obviously the same decision problem is involved in both cases, but the experimental findings show that significantly more is invested in the public project under the Give frame than under the Take frame.
- The observation that the results of experiments can be strongly influenced by the respective frame has led to the emergence of neutral frames as a standard – at least when it comes to testing general models. This means that names that could be given to an interaction or the persons involved are consciously avoided and that the description of the experiment is designed as value-free and neutral as possible.

# The Frame of the Experiment

- Let us assume that when subjects enter a laboratory and receive instructions for an experiment, they first try to understand what the experiment is about and what behavior is expected from them. The frames of the experiment then serve as an orientation aid for the subjects. What is the name of the experiment? What is the name of the activity I need to perform? What conclusions can be drawn from the type of task I am faced with here?
- Questions of this kind will occupy the subjects. It should be borne in mind that the subjects assume that the frame – i.e. the answers to their questions – was set by the experimenter. The person who wrote the instructions and designed the experiment thus provides the information that the subjects use to make sense of the experiment.
- This means that each frame – no matter how it is designed – is always associated with a potential experimenter demand effect. If one accepts this consideration, the question of whether a change of the frame impacts on the behavior can also depend on whether this alters the potential experimenter demand effect and whether this in turn has any impact.
- Of course, the behavior of subjects is not only determined by experimenter demand effects. Ideally, their influence is rather small and the effect of monetary incentives dominates the decision. Nevertheless, when designing an experiment, one should at least be aware of the potential connection between frames and experimenter demand effects.
- The second way the frame has an impact is that it can also influence the beliefs of the subjects about other subjects' behavior. This is all the more the case because the frame directly creates common knowledge.
- A third way it impacts arises because a frame can be accompanied by the activation of social norms. It is important to note that such norms can also have an influence in the real world. If a real phenomenon is to be simulated in the laboratory, a corresponding frame should therefore be included.

# Instructions and Comprehension Tests

- All the elements of an experimental design must be communicated to the subjects of the experiment. This is done in the instructions, which are either provided verbally or distributed in writing to the subjects. Two important questions are of interest here. First, how can the instructions be conveyed in such a way that it is certain that all the subjects have actually taken note of and understood them, and second, how can potentially distracting effects be eliminated?
- Ideally, instructions should be in writing and distributed as a document to the subjects. An important reason for this is that it is then certain that the subjects can look at the instructions again during the ongoing experiment if anything is unclear to them. This also rules out variations in the presentation of the instructions from session to session that can undoubtedly take place if the instructions are communicated verbally (even if arises simply through a variation in the emphasis of some words).
- However, by providing the instructions verbally, it is possible to ensure that they are common knowledge for all the subjects. In other words, the subjects know that everyone knows that everyone knows... that everyone knows what is in the instructions. It is therefore not at all unusual for the instructions to be distributed in writing, and also to be read out.
- As far as the content of the instructions is concerned, there are three points to bear in mind: (1) The description of the experiment should be as short and concise as possible. (2) The description of the experiment must be as simple and understandable as possible. (3) Instructions are the point where experimenter demand effects could be generated or norms might be triggered. This is something to be aware of, i.e. when writing the instructions it is important to remember that signals are being sent to the subjects who could possibly use them to interpret what they should do.

# Instructions and Comprehension Tests

- How should we deal with questions that the subjects still have after they have received the instructions? We recommend that questions not be asked publicly. For this reason, reading the instructions out loud should not be concluded by asking the group if anyone has a question but rather by pointing out that questions can only be asked in the strictest confidence and then answered one-on-one between the subject and the experimenter.
- Why is it not advisable to have questions asked publicly? The problem is that there is no control over what is asked. As a consequence, questions might be asked that are not about understanding the experiment, but rather about giving an indication of individual expectations or behavior or how one should behave.
- The saying that trust is good but control is better also applies to experimenters. It is therefore a good idea to check whether the subjects really have understood the experiment. Control questions are therefore important, but they also entail the risks already mentioned. They can trigger experimenter demand effects, activate norms or lead to anchoring effects.
- In any case, all the subjects should be given the *same* control questions. This means that if values are determined randomly, then this should be done once for all the subjects and not for each one individually. This ensures that the group of subjects is homogeneous in terms of subjects' previous experience.



# Interactions Between the Subjects

- In addition to the interaction with the experimenter, there may of course also be interactions between the subjects in experiments. This is obviously evident in experiments that involve strategic interactions. There are, however, different types of exchange between subjects that go beyond purely strategic interaction.
- Well-known examples of these are the reputation effects that can accompany the identification of individuals and the effects of communication. Whether such effects are possible or not depends on the experimental design. In any case, it is important to be able to assess their impact when deciding whether or not non-anonymous interactions between subjects should be possible.
- What are in fact the arguments for designing experiments in such a way that the subjects remain anonymous to each other? The most important reason is specifically to prevent reputation and communication effects. Anonymity is frequently sought due to fears of losing control over the experiment if they are permitted. As a result, however, the experimental context sometimes differs markedly from the context in which real interactions take place.
- A good example of this is experiments that deal with coordination problems. One of the workhorses in this field is what is named the minimum effort coordination game. This involves a group of players who are required to complete a task together. To this end, each individual can make a smaller or larger effort, which generates costs. The payoff to all the members of the group depends on how large the minimum effort made by an individual member of the group is. In other words, the weakest link in the chain decides. The question is what kind of equilibrium does coordination ultimately lead to when the game is repeated. It has been well known since the work of van Huyck et al. (1990) that groups of more than 4 members are generally not capable of coordinating on the payoff-dominant equilibrium.
- Is it conceivable that in the real world there are situations in which 6 or 7 people in a group have to complete a common task, the weakest group member decides on the remuneration for everyone and this happens in complete anonymity? It would be difficult to find an example of this. It is indeed a limitation to use experimental designs that are known not to exist in such a way in real situations. Instead of categorically excluding reputation and communication effects, it may therefore be a sensible strategy to consciously allow them so as not to ignore what may be an important aspect of real decision-making environments. Of course, in this case it is important to know what effects communication and reputation can have.

# Reputation Effects and Social Distance

- Does reputation really play a big role? And if so, how are reputation effects triggered? How much social interaction is necessary for subjects to start thinking about their reputation? And how does the reputation effect differ from what triggers a reduction in social distance?
- While some experiments show that reducing the social distance (e.g., showing faces, telling names etc.) do not have an effect on behavior, others, show that it does. For example, Bohnet and Frey (1999) run a double-blind dictator game, the interaction between dictator and receiver, however, was varied. In the first treatment, as a baseline treatment, the experiment was conducted anonymously. The second treatment involved one-way identification. This was achieved by having the receiver rise from his or her seat and thus be identified by the respective dictator. This simple identification had no effect. Admittedly, in view of the fact that the receivers remain completely inactive in a dictator game experiment, there is also no reputation effect.
- In light of the above, what happened in the third treatment is quite astonishing. The one-way identification described above was repeated there, but this time the receiver said their name and mentioned their favorite hobby. Although no reputation effect could occur in this arrangement either, the allocations increased significantly. It does apparently matter how familiar the other person is. Social distance is important, at least in dictator game experiments.
- The fact that social distance can influence laboratory behavior suggests that even outside the laboratory it is important how anonymously people act or how close they get to other people (see also Brosig-Koch and Heinrich 2018 whose study is based on both, laboratory and field data).
- This should be taken into account when deciding which interactions are to be permitted in the laboratory. Strict anonymity makes the experimenter's life easier because it ensures that the conditions of interaction can be well controlled. A reduced social distance is always associated with a potential loss of control. It is important to be aware, however, that anonymity can lead to certain types of behavior that do not occur with lower social distance. If the real phenomenon to be studied experimentally is not characterized by strict anonymity, experiments conducted anonymously are subject to a considerable loss of external validity.

# Communication effects

- Controlling communication: No matter how communication between the subjects is to be designed, it is important that the experimenter retains control over how the subjects interact. This involves not only the experiment itself, but also what happens before and after the experiment. It may be advisable to ensure that uncontrolled communication can be ruled out as far as possible when recruiting subjects. The same applies to the way the subjects enter the laboratory and the way they leave the laboratory after the experiment. A complete control of communication requires that all these steps are included.
- The Conflicting Objectives of Control and External Validity
- The basic problem that arises in connection with communication among subjects can be described as a conflict of objectives. If experiments are played in complete anonymity, greater control over the interaction is achieved, as effects triggered by communication can then be avoided. This facilitates the interpretation of the results and eliminates the need to isolate and identify the effects of communication. Unfortunately, completely removing communication between the people involved means that we are far removed from many real contexts in which people are active.
- On the other hand, this does not of course mean that there are never situations that are best reproduced in the laboratory using treatments that are anonymous and without communication. For example, it can be argued that the actors in (online-) markets often make decisions alone, without interaction with other people.
- The reason most experiments do not allow communication is that there is often the concern that communication can have many very different effects and that, if it is allowed, the ability to interpret the results of the experiment in a meaningful way is lost.
- On the other hand, fear of the lack of control over communication effects has led to the study of economic phenomena in speechless anonymity. Even with the best will in the world, one cannot imagine that in reality such phenomena take place even remotely under such conditions. For example, there will likely be very few negotiations in which those who negotiate never exchange words and who moreover do not know each other. Against this background, the question arises whether concerns about giving up too much control when allowing communication are really justified. It must be kept in mind that there are different forms of communication and that different techniques can be used which differ greatly in terms of control over the effects of communication.

# Forms of communication

- Communication can be used for different purposes. It can be used to transmit information that the communication partners possess. But it can also be used to gain a visual (gender, appearance, facial expression) or acoustic (dialect, emphasis) impression of the communication partners. Communication can be uni-, bi- or multidirectional. It can be face-to-face or without eye contact and messages can be spoken, written or conveyed with gestures. Even within these forms of communication there are still many possible variations. For example, face-to-face can mean that the subjects sit at the same table and talk to each other, but face-to-face can also be achieved by means of a video conference. Written messages can be communicated through a chat program or with handwritten messages.
- Further distinctions are possible. For example, the permitted communication content can be limited or unlimited. In the first case, only discussions relating to the task set in the experiment might be permitted, or the subjects may be allowed to talk about everything except the experiment. If, for example, the written form is chosen precisely because communication is permissible but reputation effects are to be excluded, it should be strictly forbidden to send messages that allow conclusions to be drawn about the sender's identity.
- Finally, the experimenter has to decide in what form and to what extent the communication should be recorded and evaluated. If, for instance, a video conference is recorded, it is possible to evaluate not only the contents of the communication, but also the gestures and facial expressions of the subjects. With the aid of suitable software, such an evaluation is now also possible by computer. Eye tracking makes it possible to determine the way in which people perceive information. This makes even unstructured face-to-face communication considerably easier to monitor.

# Communication Effects

- The analysis of communication effects should take place against the background of the economic evaluation of communication. The focus here is on the game-theoretical concept of “cheap talk”. In general, this means communication that does not affect players’ payoffs. This form of communication can have behavioral effects if the interests of the players are sufficiently similar. However, if players have conflicting interests – such as in the prisoner’s dilemma – this form of communication should not influence their actions.
- Cheap talk is a strategic interaction when it is not possible for players to check the truth of the information they receive from other players and when it is possible to lie without incurring costs. From a game-theoretical point of view, experiments in which players have conflicting interests and communication between each other are completely harmless – at least if this communication is merely cheap talk. Since cheap talk is not supposed to change behavior here, it can be ignored.
- However, when cheap talk can still be assumed depends on whether or not the liar incurs costs. Since psychological causes for such costs – which cannot be directly observed – are also possible, it is therefore conceivable that communication is not “cheap” at all, although at first sight it appears to be so. Thus, also from a theoretical point of view, it cannot be ruled out that communication may have an effect in a great many contexts and games.
- What do the experimental findings look like? At which points is it relatively easy to imagine that communication between the subjects has an effect? The first thing that comes to mind is the experiments that deal with the coordination problem, such as the minimum effort coordination game. It really does not make any sense to carry out such experiments anonymously and without the possibility of communication, because such situations are hardly likely to be found in the real world. The reason behind this was of course the expectation that the coordination problem would be more or less resolved if those who were faced with it could communicate with each other.
- In fact, according to the game-theoretical prediction, communication in the minimum effort coordination game can also have behavioral effects due to the common interest of the players to achieve the payoff dominant equilibrium. For example, if players mutually promise to put in their best effort, there is no incentive for players to falsely state the level of effort they intend to play. Lying is not a rational strategy in this game. This lends a high degree of credibility to the pronouncements, which in turn enables the players to use communication to make the payoff-dominant solution a kind of focal point that everyone is guided by.

# Communication Effects

- It is not only in pure coordination games that it is advantageous to be able to mutually coordinate behavior. Even when it comes to striking cartel agreements, there is reason to believe that being able to consult with each other could have an impact on the formation and stability of such agreements. In this case, however, - assuming strict selfishness - we are dealing specifically with a game with a dilemma, which means that the interests of the players are not similar. Although everyone has an interest in the others abiding by the agreement, the individual would prefer to deviate from it.
- Fonseca and Normann (2008) examined whether this is actually the case. They conducted an experiment in which 2, 4, 6 or 8 players were in Bertrand competition. Each treatment was played in two variations, one without communication and one with the possibility to consult with each other via a chat program for 1 minute. Although this is not exactly an excessive form of communication and although theoretically it should not trigger any behavioral effects, it did have a clear impact. With chat, the prices that companies set were higher than those without chat and corporate profits increased when communication was possible.
- Communication, as the explanations so far have shown, has a very strong effect in many experiments. It can facilitate cooperation and makes collusion more likely. It can lead to fair negotiated solutions and can increase trust and trustworthiness in the trust game. The question is, why is it that it leads to these effects? When deciding whether and in what form communication should be allowed in an experiment, it is helpful to know the channels through which communication can influence behavior. It is not clear whether all channels really are known and whether we already have a comprehensive understanding of the effect of communication, but some statements can be made which can claim some plausibility and for which experimental evidence is available.

# Possible Causes of Communication Effects

- How and why communication works depends to a large extent on the context in which it takes place and on the form of communication. The question of whether or not eye contact is associated with verbal communication has turned out to be critical for the sustainable effect of communication. The combination of language and visual identification is obviously important. The simple identification of the other person does not in itself make much difference, but the face-to-face exchange of information results in marked changes in behavior. It should come as no surprise that this form of communication plays a special role.
- For a very long time, face-to-face communication was the only form of communication. Evolutionarily, therefore, it may have played an important role. But it is also of paramount importance in the individual socialization process of each person. Long before learning to use other communication channels, people meet their closest caregivers almost exclusively face to face. These are of course pure plausibility considerations, but they are consistent with the experimental evidence we have reported on.
- A reliable explanation of communication effects can be linked to two points. Either reputation effects that are caused by communication change the strategic situation – and thus the equilibrium – or the personal encounter with the other players changes the attitude towards them or provides additional information that leads to a different perception of the decision situation.
- The experimental evidence suggests that reputation effects alone may play a rather minor role. The experiments have shown that it is not enough for subjects to be able to identify one another visually in order to trigger behavioral changes. This speaks in favor of the second point that the perception of the decision-making situation changes when communication takes place.
- A possible explanation for how this could happen can be found on the basis of “psychological game theory”, Charness and Dufwenberg (2006) show that the effect of communication in dilemma situations can be explained by guilt aversion. The theory can be outlined as follows: Two players, A and B, are in a situation where player B can receive something from player A (a transfer payment, a contribution as part of cooperation, or similar). A assumes that B has a certain belief about A concerning this. On top of this, A himself forms a belief, i.e. A can either fulfill or disappoint the expectation that he believes B to have of him. In the latter case, the anticipated disappointment of B resulting from this may lead to a feeling of guilt in A. If people do not like feeling guilty, fulfilling the beliefs serves to ward off this feeling. The decisive factor here is that communication, and in particular face-to-face communication, can change A’s beliefs regarding B’s beliefs. This is usually done by way of promises made during communication. If, after a conversation with B, A believes B believes he will receive more from A than A had assumed before the conversation, the pressure on A to increase the amount he gives to B in order to avoid feeling guilty increases.

# Decisions Made by the Subjects

- Laboratory experiments are about presenting subjects with questions and observing their decisions under controlled conditions. In a sense, experimenters direct questions to the subjects, who answer them in making their decisions.
- But how should these questions be formulated? And in which form should the answers be collected? There is no one definitive answer to these two methodological questions, as there are different approaches to take and methods to use and all have their advantages and disadvantages.
- Therefore, the experimenters first have to make a decision before the subjects do: Which experimental design is the best for our specific experiment? An answer can only be found if the research question on which the experiment is based is known and if the hypotheses for the experiment have been established.
- Both, the formulation of the research question and the establishment of hypotheses are, therefore, important first steps on the way to a suitable experimental design. Following slides provide an overview of the options available for eliciting decisions.



# Strategy Method Versus Direct Response

- It is generally easy to determine the elicitation method to use in experiments involving the decision-making behavior of individual subjects without the occurrence of any strategic interaction. The subjects are presented with a specific decision problem, i.e. they have to make a choice, and it is this choice that is observed. The matter can become much more complex if strategic interactions arise in the experiment. It is, in the first instance, irrelevant whether the game played by the subjects takes place simultaneously or sequentially. For better understanding, however, it is simpler to assume a sequential game.
- Direct elicitation (“hot”) vs. with the strategy method (“cold”).
- The normal case is that the players make their moves in the order specified, with the second mover responding to the move made by the first mover, the third mover reacting to that of the second mover, etc. The players thus provide a direct response to the action of the mover before. This method of eliciting the responses is simple and easy to understand. From the point of view of the experimenter, however, it can have a considerable drawback.
- Let us take the simplest sequential game imaginable. Two players each choose between two possible alternatives. In this case, there are four possible outcomes of the game. Each individual decision that is observed, however, only provides information about one of the four possible paths on which the game tree can be traversed. Suppose the first mover has a choice between alternatives a and b. If the first mover (for whatever reason) has a preference for a, and chooses this strategy in nine out of ten cases, it becomes quite difficult and expensive to collect enough observations in the subgame following b.
- The strategy method, which essentially goes back to an article by Selten (1967), offers an elegant solution to this problem. Instead of the second mover being presented with the decision of the first mover, he is required to specify a complete strategy. In our simple example, he has to indicate what he will do at the two decision-nodes he can reach. In other words, he must indicate how he will respond in both cases, i.e. if first mover plays a and if he plays b. The result of the game is obtained by combining the move chosen by the first mover with the corresponding response from the strategy of the second mover. In this way, the experimenter elicits information about behavior throughout the game.

# Experiments with Real Effort

- Economic experiments almost always involve decisions in which costs play a role, whether it is a case of the subjects being faced with an allocation task in which every amount they give is at the expense of their payoff, purchasing goods or making a contribution to the production of goods. Occasionally, the work efforts that are exerted to fulfill a task are also represented by appropriately designed cost functions (for example, in the minimum effort coordination game). A two-stage procedure is usually used to implement costs in the laboratory. The first stage consists of giving the subjects an income in the form of an initial endowment (house money). This income can then be used to cover the costs incurred. In the second stage, the costs are specified in the form of a mathematical function, with there being considerable room for creativity. For example, the cost function can be convex to represent that it becomes increasingly difficult to exert the effort.
- Inducing costs in this way has considerable advantages, especially in view of the fact that the experimenter retains complete control. Since the costs are part of the payoff function, it is indisputable to what extent they are actually incurred. However, this high degree of control comes at a price. People may treat the money they are given differently from the money they have earned from work. It is therefore not entirely unproblematic to first give subjects money that they can then use to cover costs.
- An alternative to issuing house money is to have the subjects work for the money they receive by introducing real effort. This increases external validity and avoids the house money effect, but has the disadvantage that the control over costs is lost. If subjects are allowed to “work” in order to impose costs on them, the actual level of costs that the subjects incur depends on the burden of the work they have to bear – and that cannot be observed! The question is, under which conditions a real effort design is appropriate and, above all, how it can be designed in concrete terms.
- An important requirement is that it be structured in such a way that it can be assumed that at least at the beginning of the experiment all the subjects are equally good at achieving this performance. Therefore, no prior knowledge that may exist to varying degrees should be required and personal aptitude should not play an important role. It is also clear that the task should be easy to explain so that the subjects understand what is involved. Furthermore, the work outcome should be easily and reliably measurable and allow a comparison between the subjects. Finally, the task should be designed in such a way that possible learning effects are minimized and quantifiable, so that these effects can be corrected if necessary.

# Within- Versus Between-Subject Design

- At the core of experimental research stands the comparison of different experimental treatments under controlled conditions. An experiment that consists of only one treatment makes relatively little sense. It is almost always a case of subjects making decisions under different conditions, with the treatments that are being compared as far as possible differing in only one parameter, thus enabling conclusions in relation to causality to be made.
- A fundamental issue of experimental design in this regard is whether each individual subject participates in a number of different treatments or whether every treatment involves different subjects, with each subject participating in only one treatment. The first case is described as a “within-subject design”, since the comparison takes place within one and the same subject, while the latter case is called “between-subject design” due to the comparison between the subjects.
- Within-Subject design
  - Advantages: the number of observations per subject is greater when each subject participates in several treatments than when new subjects are invited, the internal validity of the experiment does not require successful randomization to have been carried out, closer proximity to theory - higher external validity.
  - Disadvantages: there is no avoiding that dependencies arise between the individual observations in the different treatments (can be solved by taking into account the order effects, and by panel data analysis), presenting the subjects with different treatments can lead to an experimenter demand effect
- Between-Subject design
  - Advantages: easy to handle - the only condition to be met is that the subjects are randomly assigned to the different treatment. the statistical analysis of between data is easier than within data, as it is not necessary to correct for dependencies between data elements. Also, between designs tend to lead to conservative results as compared to within designs, therefore there is a relatively high certainty that this finding is revealing a causality.
  - Disadvantages: considerably more resources (time, money, subjects) may be required to obtain statistically meaningful data than with the within design approach. In other words, with the same use of resources, less statistical “power” is likely to be achieved with a between design than with a within design. Moreover, the external validity is not as direct as with a within design.

# The Repetition of Games

- There are only very few decisions of economic interest that we take once only and are never faced with again. Normally we have to make decisions again and again. In fact, we may even make some of them very frequently. In a certain sense, this is a good thing because it gives us the opportunity to learn and adapt our behavior to experience. This is also the reason why many games are played repeatedly, meaning that the subjects make the same decision several times within one experiment. The methodological implications that this has depend on how the repetitions are designed.
- For example, for experiments in which strategic interactions occur, it makes a significant difference whether this interaction takes the form of a repeated “one-shot” game – i.e. with a new partner in each round – or whether it is a repeated interaction with one and the same partner. It is possible – although rarely done in practice – to repeat experimental sessions with the same subjects. Here, too, a few things have to be taken into account so that the data obtained can still be meaningfully interpreted.
- Repetition Within a Session
- The majority of games tested in economic experiments are played over several rounds, i.e. the respective game is played repeatedly with the same subjects within one session. The main reason for these repetitions is to give the players the opportunity to gain experience and learn the game. To a lesser extent, it is also to investigate whether and how behavior changes when one and the same task has to be solved repeatedly.
- Repeating games within a session can create different learning effects. The experimenter determines which are possible, at least in part, through the design of the experiment. For example, if the same partners always play the game together, the learning opportunities that arise are different from those that occur when playing with new subjects in each round. If the experiment is to be designed in such a way that one-shot interactions are to be repeated, then a round-robin design should be chosen because this is the only way to avoid direct and indirect reputation effects.
- Which learning effects occur also depends on the complexity and type of the game. In more complex games, repetition can lead to a better understanding of the game.
- In repeated games, it is important whether the test subjects know how often they play or whether they are unaware of it. Games played with random termination can be used as a (more or less close) substitute for unending games

# The Repetition of Sessions

- It is very common for games to be repeated within a session. This is probably the case in the majority of experiments. It is very rare, however, that entire sessions are repeated identically with the same subjects.
- One reason for this may well be the fact that it is not possible to control what the subjects do between sessions, thus implying a loss of control.
- Another problem with repeated sessions is unreliable subjects who do not attend all the sessions. Subjects failing to show up is not only annoying, but also reduces the interpretability of the experimental results because it cannot be ruled out that a selection process is associated with their absence. For example, it could be precisely the subjects who had certain kinds of experience in previous sessions who are missing in later sessions. A very effective means against such absences is to postpone payments to the end of the series. The threat of going away empty-handed if one does not show up for all the sessions is quite credible and should not fail to have the desired effect.
- On the other hand, repetition may lead to an increase in external validity and allow mature behavior to be observed in the laboratory.
- The loss of control can, in particular, be minimized by ensuring that the probability of contact between the subjects is as low as possible. This is achieved first and foremost by maintaining strict anonymity during the experiment.

# The Reproducibility of Experiments

- Scientific research aims to make general statements about causal relationships whose validity can be verified intersubjectively. An experiment conducted in a particular laboratory at a particular time with particular subjects cannot be the basis for such a statement. Its results are in the first instance no more than a single observation. If a causal relationship is established, it applies specifically to this experiment, and it cannot easily be generalized. Experimental findings become usable – to put it more precisely – actually only when they have been proven several times and it has been shown that they apply irrespective of time and place.
- In the final analysis, however, it is not the aim of experimental research to produce “true” statements in the above sense. Rather, it is about creating empirical, experimental evidence. This means gaining insight into which causal relationships are likely to be encountered under which circumstances. This is not possible with a single experiment, but requires the experience that observations can be reproduced relatively reliably. This has some methodological implications, also for the design of experiments. The most important implication is that experiments must always be designed in such a way that they are reproducible.
- However, it is not enough to simply create a design that allows an experiment to be reproduced. So that a replication can actually be conducted, this design and all its elements must also be well documented. Comprehensive documentation must guarantee that the experiment can be conducted identically by other people at a different location. This means that details of all the auxiliary materials used must be supplied. Not only the instructions that the subjects received, for instance, but also the software used in the experiment should be available to those who want to reproduce it.
- In addition, the procedure of the experiment must be documented very precisely. This includes, for example, the way in which the subjects were invited and received in the laboratory, whether and in what form they had contact with each other, how the instructions were distributed, whether they were read aloud and how questions of understanding were dealt with. Every single detail could be important. Of course, the raw data collected in an experiment are also among the things that have to be documented.
- Nevertheless, it remains a fundamental problem of experimental economic research that there is a lack of incentive to carry out the important task of reproducing experiments - replication papers are usually not being published in important journals.