

Incentives

- Today we will discuss the impact of incentives on people's behavior in general. The emphasis will be in how a certain incentive (a wage, a fine...) can affect the "norms" that surround transactions.
- Gneezy and Rustichini: "Pay enough or don't pay at all", Quarterly Journal of Economics.
- Can rewards have detrimental effects on performance? Does the introduction of monetary exchange change the "contract"?
- Think about blood donation. Should we provide monetary incentives to provide blood? Wouldn't that affect who gave blood (and hence the quality of blood) and would it necessarily increase the supply of blood?

Experiment 1:

- 160 students from the university of Haifa are asked to answer 50 questions taken from an IQ test. All participants receive 60 NIS show up fee. The questions are ones that involve reasoning and computation as opposed to just general knowledge.

4 different treatments that vary in how much additional money subjects receive for each question solved.

- No additional incentives.
- 10 cents of a NIS per correct question.
- 1 NIS per correct question.
- 2 NIS per correct question.

TABLE I
SUMMARY STATISTICS FOR THE IQ EXPERIMENT, FOR THE DIFFERENT TREATMENTS
The Lower Fraction is the Fraction of Subjects Who Gave a Number of Correct
Answers Less than 16

	No payment	10 cents	NIS 1	NIS 3
Average	28.4	23.07	34.7	34.1
Standard deviation	13.92	14.72	8.88	9.42
Median	31	26	37	37
Average top 20	39	34.9	42.35	41.6
Standard dev. top 20	5.25	6.79	3.63	4.18
Average bottom 20	17.8	11.25	27.05	26.6
Standard dev. top 20	11.56	10.22	5.07	6.82
20th quantile	40	35	44	43
80th quantile	20	0	26	25
Lower fraction	15%	27.5%	0%	0%

TABLE II
MANN-WHITNEY U TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS
OF MEDIANS OF CORRECT ANSWERS BY TREATMENT

	No payment	10 cents	NIS 1
10 cents	.0875	—	—
NIS 1	.0687	.0004	—
NIS 3	.0708	.0006	.6964*

(Prob. $> |z|$, where z is the test statistic). An asterisk indicates that for that comparison we cannot reject (at a .9 level of significance) the hypothesis that the two samples come from the same distribution.

Similar results when looking at 20 best and 20 worst observations

How about a real effort task, in which the effort of participants is directed towards something “productive” as opposed to “just” solving IQ questions for the experimenter?

Experiment 2: the Donation Experiment.

In Israel a few “donation days” take place every year. Each of these days is devoted to a society that collects donations from the public for some purpose, such as cancer research, disabled children, etc. High-school students go from door to door to collect the donations.

Students are put in pairs. Each pair receives a certain number of coupons, which serve as receipts for the donors. The amount collected by each pair on the donation day depends mostly on the effort invested: the more houses they visit, the more money they collect. This is especially true because the students do not have to “sell” the donation, since most people are already familiar with it from television announcements and advertisements.

180 high school students around the age of 16, each pair received coupons for NIS 500.

- Group 1: Amounts collected by group will be made public.
- Group 2: Each group receives 1% of the amount collected.
- Group 3: Each pair receives 10% of the amount collected.

In each case: Payment is made out of funds from the experimenter, not the money collected (and that is known by participants in advance.)

TABLE IV
SUMMARY STATISTICS FOR THE DONATION EXPERIMENT,
FOR THE DIFFERENT TREATMENTS

	No payment	1 percent	10 percent
Average	238.6	153.6	219.3
Standard deviation	165.77	143.15	158.09
Median	200	150	180
Average top 20	375.33	272	348
Standard deviation top 20	111.92	98.64	110.46
Average bottom 20	102	35.33	90.66
Standard deviation bottom 20	66.13	52.08	63.97
20th quantile	100	0	50
80th quantile	450	250	400

TABLE V
MANN-WHITNEY *U* TESTS BASED ON RANKS WITH PAIRWISE COMPARISONS
OF MEDIANS OF AMOUNTS OF MONEY COLLECTED BY TREATMENT

	No payment	1 percent
1 percent	.0977	—
10 percent	.7054*	.0515

How do participants perceive the effects of monetary incentives?

- The IQ Experiment

53 students will be “principals” and will be matched with one player, the “agent” from the IQ experiment.

They receive 1 NIS for each correct answer of the agent.

The principal can decide whether the agent should be paid 0 or 10 cents of NIS for every correct answer.

The principals are told that the agent would know in advance how much he would get paid, but not that there was a choice about how much he'll be paid, nor will he know an agent exists in the first place.

46 subjects (that is 87%) chose to pay 10 cents for every correct answer.

- The Donation Experiment

Principals were told they would be matched with one pair that had already collected money, and would be paid 5% of what that pair collected.

The principals had to decide whether the pair should be chosen from the group that did not receive any payoff or from the group that received 1 percent of what they have collected. The payment to the agent was made out of the 5 percent.

The results confirmed the previous test.

Out of the 25 participants, 19 (that is, a proportion of 76 percent) preferred to be matched with an agent who was paid 1 percent of the amount he collected.

- One possible explanation is that low incentives (like 10 cents of a NIS) are insulting..
- Another possible explanation, is that the introduction of monetary compensations “changes” the contract.
- How would we test among these hypotheses, and would this nonmonotonic effect also exist when we consider negative instead of positive incentives?

Gneezy and Rustichini: “A fine is a Price”, Journal of Legal Studies, 2000.

10 private day care centers in the city of Haifa. Each one can hold at most 35 children from the age of 1 to 4. The fee for each child is NIS 1400 per month. The contract signed at the beginning of the year states that the day care center operates between 7:30 and 16:00.

In particular, before the beginning of the study, there was no fine for coming late. When parents did not come on time, one of the teachers had to wait with the children concerned. Teachers would rotate in this task, which is considered part of the job of a teacher, a fact that is clearly explained when a teacher is hired. Parents rarely came after 16:30.

The overall period of the study was 20 weeks. In the first 4 weeks we simply recorded the number of parents who arrived late each week. At the beginning of the fifth week, we introduced a fine in six of the 10 day-care centers, which had been selected randomly. The announcement of the fine was made with a note posted on the bulletin board of the day-care center.

Parents tend to look at this board every day, since important announcements are posted there. The announcement specified that the fine would be NIS 10 for a delay of 10 minutes or more.

Announcement: Fine for Coming Late

As you all know, the official closing time of the day-care center is 1600 every day. Since some parents have been coming late, we (with the approval of the “Authority for Private Day-Care Centers in Israel”) have decided to impose a fine on parents who come late to pick up their children.

As of next Sunday a fine of NIS 10 will be charged every time a child is collected after 1610. This fine will be calculated monthly, and it is to be paid together with the regular monthly payment.

Sincerely,

The manager of the day-care center

At the beginning of the seventeenth week, the fine was removed with no explanation. Notice of the cancellation was posted on the board. If parents asked why the fines were removed, the principals were instructed to reply that the fine had been a trial for a limited time and that the results of this trial were now being evaluated.

TABLE 1
 NUMBER OF LATE-COMING PARENTS PER WEEK ACCORDING TO DAY-CARE CENTERS

CENTER	No. OF CHILDREN	WEEK																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Treatment group:																					
1	37	8	8	7	6	8	9	9	12	13	13	15	13	14	16	14	15	16	13	15	17
2	35	6	7	3	5	2	11	14	9	16	12	10	14	14	16	12	17	14	10	14	15
3	35	8	9	8	9	3	5	15	18	16	14	20	18	25	22	27	19	20	23	23	22
4	34	10	3	14	9	6	24	8	22	22	19	25	18	23	22	24	17	15	23	25	18
5	33	13	12	9	13	15	10	27	28	35	10	24	32	29	29	26	31	26	35	29	28
6	28	5	8	7	5	5	9	12	14	19	17	14	13	10	15	14	16	6	12	17	13
Control group:																					
7	35	7	10	12	6	4	13	7	8	5	12	3	5	6	13	7	4	7	10	4	6
8	34	12	9	14	18	10	11	6	15	14	13	7	12	9	9	17	8	5	11	8	13
9	34	3	4	9	3	3	5	9	5	2	7	6	6	9	4	9	2	3	8	3	5
10	32	15	13	13	12	10	9	15	15	15	10	17	12	13	11	14	17	12	9	15	13

NOTE.—The columns under “week” report the number of parents who came late.

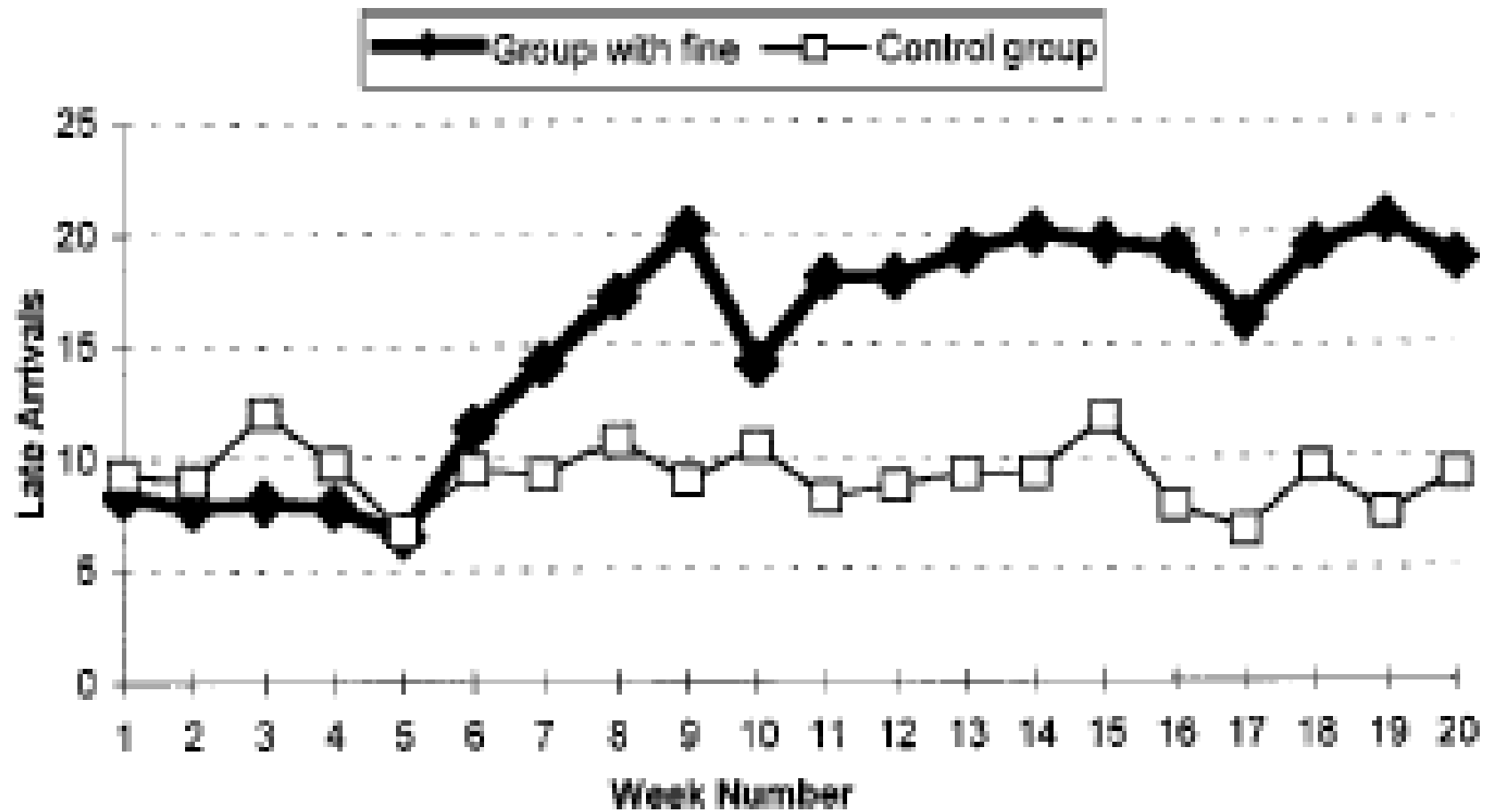


FIGURE 1.—Average number of late-coming parents, per week

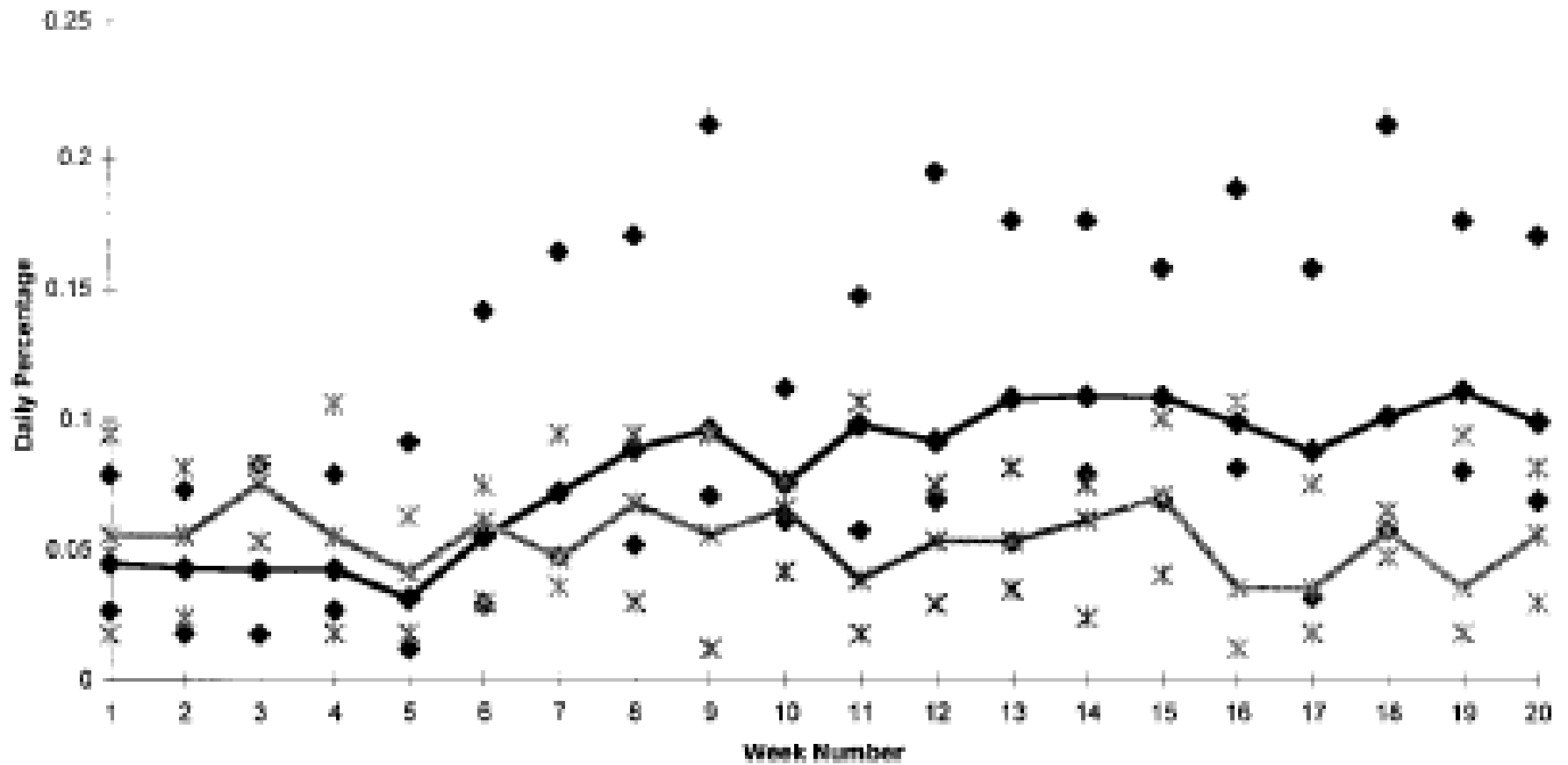


FIGURE 2.—Median value of delay for the test (dark line) and the control (light line) groups. The diamonds and the crosses indicate the extreme values for the test and the control groups, respectively.

- So, we see an effect from introducing incentives compared to no incentives.
- In all cases the incentives change the “contract”, the situations seem some in which introducing small incentives (the possibility of small fines or rewards) reduces proposals (effort), but proposals (effort) seem to increase as the size of the incentive increases.
- Would similar effects work when you think about self-motivation?
- What about the negative affect of very large incentives, such as choking under pressure? (Ask your friends that just went on the job market about the potential of that problem...)

Ariely, D., U. Gneezy, G. Loewenstein and N. Mazar (2008)
“Large Stakes and Big Mistakes.” Forthcoming in the
Review of Economic Studies.

- To test for potential negative effects of very high stakes: go to developing countries: e.g. rural India: Average monthly expenditure: ~ 500 rupies (~\$10).
- Participants will be paid either 4, 40 or 400 rupies for each achieved goal.
- Furthermore, the design should not allow for the possibility of an income effect, so, the reward or payment is a function of achieving certain goals, rather than a linear function of the amount of correct solutions to a certain task.

The Experiment in Rural India:

87 people are randomly assigned to 4, 40 or 400 rupies condition. They played 7 different games in random order. For each game they receive the full amount if they reach the “very good” goal, and half the amount if they reach the “good” goal.

The 7 tasks were:

- Packing quarters
- Simon
- Recall last 3 digits
- Labyrinth (motor skill task)
- Dart Ball (motor skill task)
- Space Force (motor skill task)
- Lying about a coin flip (someone had to guess the result of a coin flip the participant knows wrongly)

Performance in the different games presented a raw scores, % of individuals who reached the ‘good’ and the “very good” performance levels

	Mean score			% “Good”(*)			% “Very good”		
	Low	Mid	High	Low	Mid	Hig	Low	Mid	Hig
Packing Quarters	202.0 (65.4)	185.7 (70.5)	235.9 (12.9)	28.6	43.3	10.3	25.0	33.3	0
Simon	6.5 (2.1)	6.3 (1.4)	5.2 (1.4)	64.2	76.7	44.8	32.1	16.7	3.5
last 3- digits	4.9 (2.7)	5.5 (2.8)	4.6 (2.4)	64.3	73.3	58.6	42.9	36.7	20.7
Labyrinth	5.9 (2.5)	4.6 (1.8)	4.1 (1.8)	64.3	50.0	27.6	21.4	3.3	3.5
Dart Ball	2.8 (2.0)	3.6 (2.6)	2.9 (1.7)	25.0	40.0	37.9	10.7	23.3	6.9
Space Force	1.8 (2.1)	1.8 (3.1)	1.2 (1.5)	25.0	23.3	17.2	21.4	20.0	3.5
Lying on coin flip	5.4 (1.8)	5.3 (1.7)	4.7 (1.4)	89.3	80.0	86.2	53.6	53.3	17.2

(*) “Good” contains the percentage of people who reached at least the “good” level.

Average number of games in which participants reached the “very good” level:

Experiment 2

Aggregated results

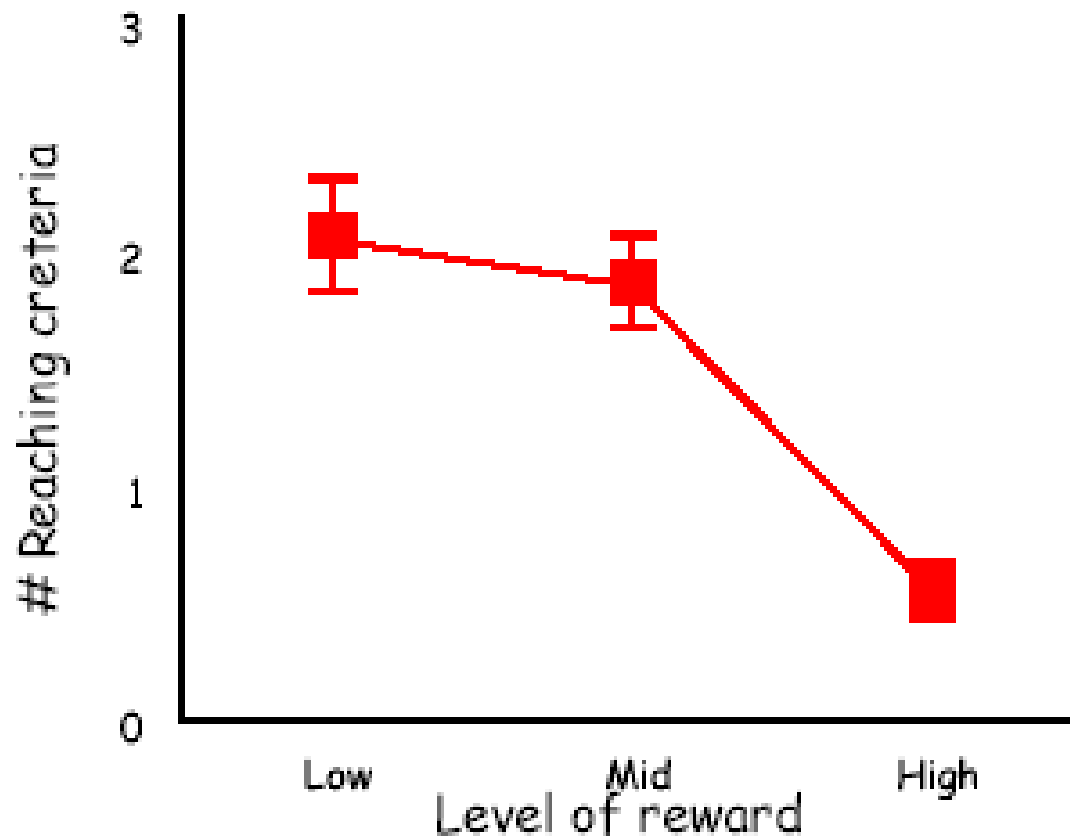
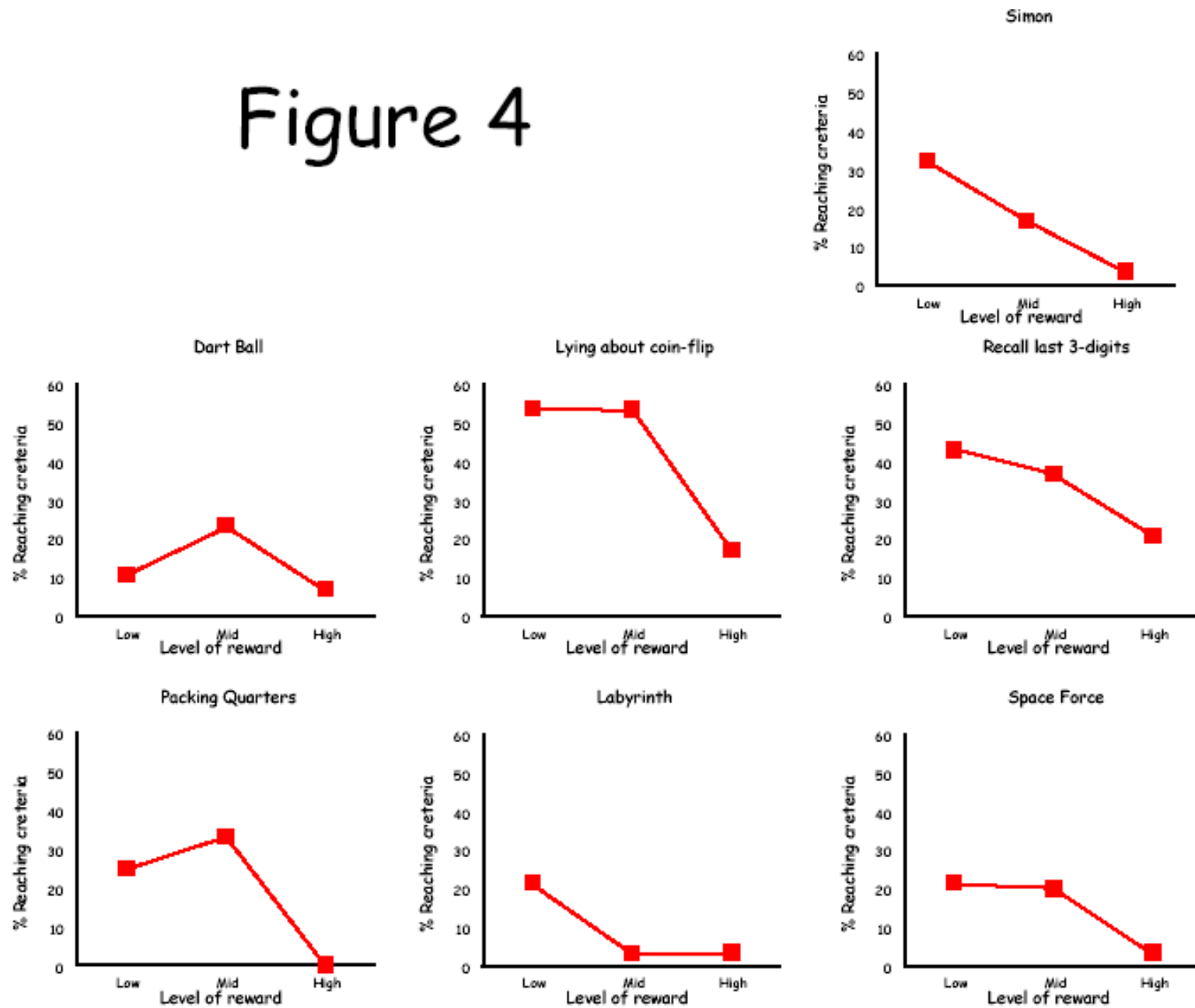


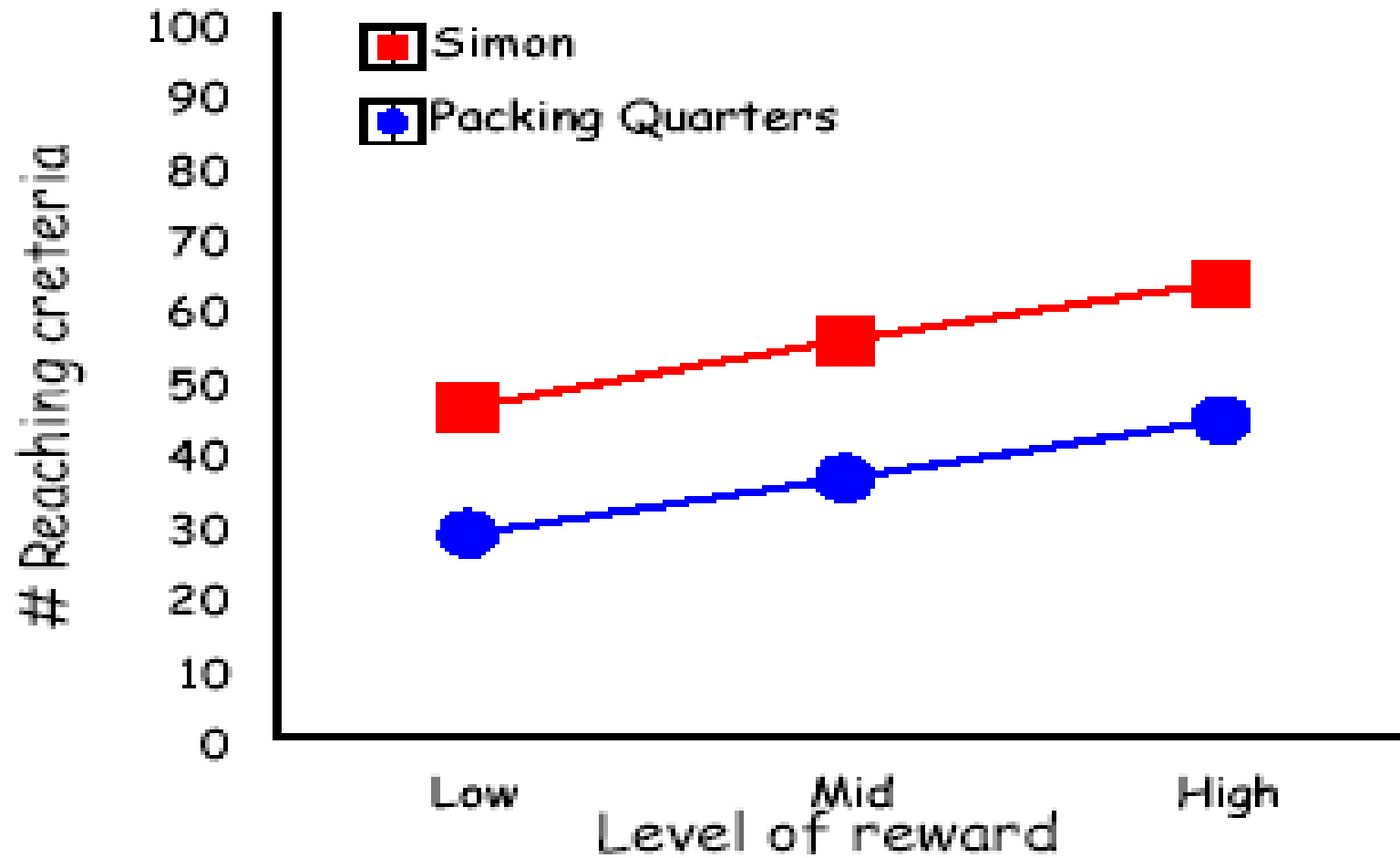
Figure 4



- First the authors ask MIT students to predict the performance as a function of the payment for the games:
- “Packing quarters”: game in which participants were asked to fit 9 metal pieces of half circles into a black wooden frame in not more than 120 seconds.
- "Simon" was described as an electronic game that required memory and repetitions (almost all subjects were very familiar with this game). It was explained that "Simon" flashes a sequence of colored lights accompanied by the light-specific sounds and that the task of the real participant is to repeat the sequence by pushing the corresponding light-buttons in the same order and manage to make at least one repetition of 8 consecutive lights within 10 trials.

Experiment 1

Predicted results



Talking about incentives

There is a whole other literature thinking about the effects of incentives.

Econ experiments: pay participants to obtain information: Does it really make a difference?

It is not only the amount of money earned that matters, but the amount that would be foregone when the answers isn't correct: There was a whole controversy on that in the auction literature.

Does the size of payments make a difference?