Is Consistency Procedure Invariant

Vered Kurtz[1] Dino Levy[2] Dotan Persitz[1]

[1]Faculty of Management
Tel Aviv University

[2]Faculty of Management and Sagol School of Neuroscience
Tel Aviv University

November 2017

Procedure Invariance

- The Procedure Invariance requirement: Recovered preferences (or heuristics) should be independent of the elicitation method.
- Necessary condition for general external validity of experiments.

Choices from Linear Budget Sets

- Choice from linear budget set is fundamental in Economics.
- Samuelson (1938), Afriat (1967) and Varian (1982) provide a formal nonparametric theory of revealed preferences in this context.
- Laboratory experiments where subjects are asked to make choices from multiple budget sets, provide relatively large individual level data sets natural for the application of the theory of revealed preferences.

Three Experimental Designs

- Three setups are used in those studies.
- ***The Textual methodology*** - subjects are faced with a sentence that describes a budget set and are asked to plug in their preferred bundle.
- ***The Graphical methodology*** - subjects are required to choose their preferred bundle from a visually presented budget set.
- ***The Discrete methodology*** - subjects are asked to choose from a small set of images (or sentences) that represent the available bundles.
- These methodologies are used to investigate:
  - Preferences over goods (bundles of various food items)
  - Risk preferences (bundles of Arrow securities).
  - Other-regarding preferences (bundles of Dictator game outcomes).
  - Time preferences (bundles of payments at different dates).

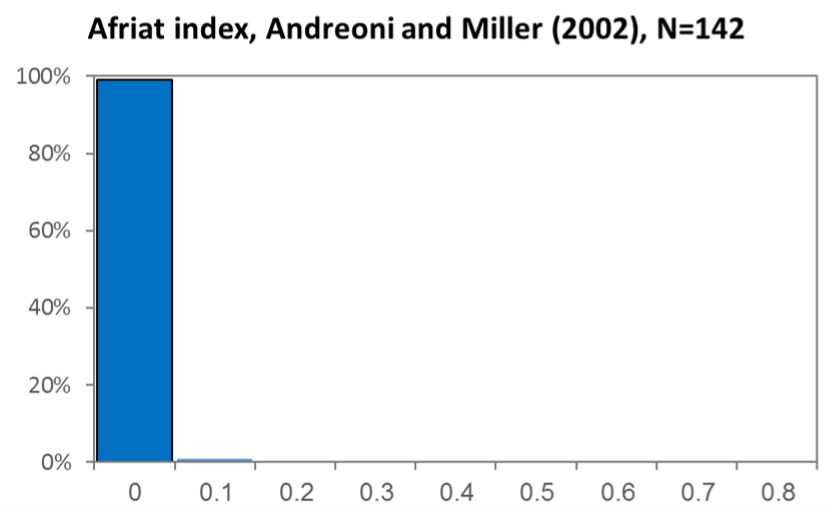
Experimental Literature

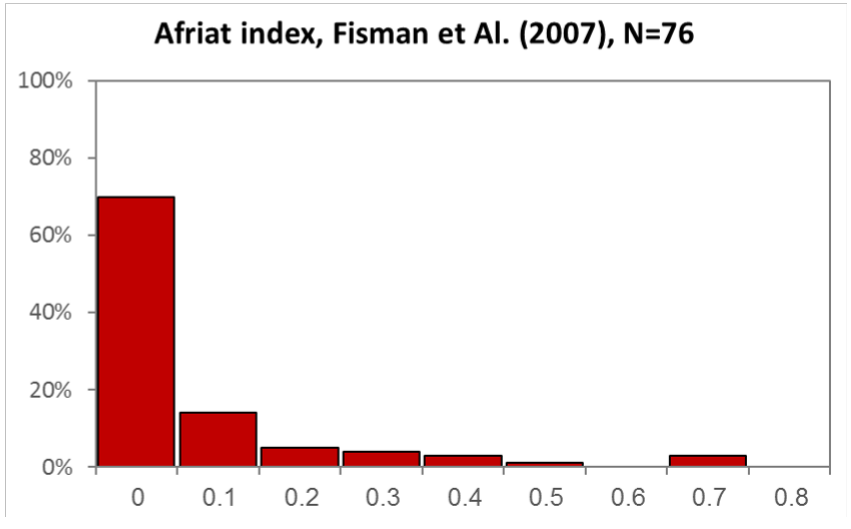
Method	Article	Preferences	Subjects	Trials	Max Slope	GARP Passing Rate	Av. CCEI
Textual	Andreoni and Vesterlund (2001); Andreoni and Miller (2002)	Giving	142 (176)	8 (11)	3 (4)	90.8% (89.8%)	0.997 (0.998)
	Eckel and Grossman (2003)	Giving	181	12	2	N/A	N/A
	Harrison and Johnson (2006)	Giving	173	10	4	N/A	N/A
	Andreoni (2007) ($n = 1$)	Giving	120	5	4	96.7%	0.996
	Dickinson (2009)	Giving	152	11	4	85% (WARP)	N/A
	Dawes et al. (2011)	Giving	234	5	3	94% (WARP)	N/A
	Visser and Roelofs (2011)	Giving	106	5	3	N/A	N/A
	Rigdon and Levine (2011), WP	Giving	189	8	3	N/A	N/A
	Dawes et al. (2012)	Giving	20	5	3	N/A	N/A
	Andreoni and Sprenger (2012a)	Time	86	5*9	2	N/A	N/A
	Andreoni and Sprenger (2012b)	Time	80	7*2	1.43	N/A	N/A
	Jakiela (2013)	Giving	144	10-12	3	N/A	N/A
	Korenok et al. (2013)	Giving	178	18	4	66% (MI)	0.979
	Kuhn et al. (2014), WP	Time	143	5*11	2	N/A	N/A
	Ashton (2015), WP	Time	149	5*9	1.5	N/A	N/A
	Porter and Adams (2015)	Giving	190	11	4	88.4%-90.5%	0.990-0.995
	Hong et al. (2015)	Social	144	20	10	56.9%	75% >0.9
	Engle-Warnick and Mishagina (2016), WP	Giving	156	20	10	30.1% (WARP)	Approx. 0.929
	Schumacher et al. (2017)	Giving	581	3	2	N/A	N/A
	Carvalho et al. (2016)	Time	1191	4*3	1.03	N/A	N/A
Graphical	Choi et al. (2007a) ($p = \frac{1}{2}$)	Risk	47	50	unbounded	25.5%	0.934
	Fisman et al. (2007) (two person)	Giving	76	50	unbounded	10.5%	0.892
	Hammond and Traub (2012), WP	Risk	41	16-48	unbounded	< 48.7%	N/A
	Choi et al. (2014)	Risk	1182	25	unbounded	22.8%	0.881
	Chow (2014)	Risk	180	20	7	N/A	0.74 ; 0.90
	Fisman et al. (2015a)	Giving	72	50	unbounded	N/A	0.944
	Fisman et al. (2015b)	Giving	208 ; 309	50	unbounded	N/A	0.95 ; 0.86
	Cappelen et al. (2015), WP	Risk	126 ; 110 ; 106	50	unbounded	23.8% ; 10%-25%, 10.4%	0.95 ; 0.856 ; 0.869
	Augenblick et al. (2015)	Time	80	5	2	N/A	N/A
	Halevy et al. (2017)	Risk	203	22	4	45.3%	0.979
	Müller et al. (2017), WP	Giving	116	50	unbounded	N/A	0.96
	Carvalho et al. (2016)	Risk	3110	25	unbounded	83%	0.81
Discrete textual	Castillo and Cross (2008)	Giving	112	4	3	N/A	N/A
	Banerjee and Murphy (2011)	Goods	69	10	5	53.6% (WARP)	N/A
	Andreoni et al. (2015)	Time	86	6*4	2.22	N/A	N/A
	Owens (2016)	Giving	N/A	50	10	N/A	N/A
	Giné et al. (2017)	Time	2142	5*2	2	N/A	N/A
Discrete Visual	Harbaugh and Krause (2000)	Giving	40	11	4	55%	0.87
	Harbaugh et al. (2001)	Goods	31 ; 42 ; 55	11	4	26% ; 62% ; 65%	0.93 ; 0.96 ; 0.94
	Camille et al. (2011)	Goods	9 ; 22	11	4	11.1% ; 68%	0.9 ; 0.95
	Bruyneel et al. (2012), WP	Goods	39 ; 31 ; 30	9	9	31% ; 48% ; 53%	0.604 ; 0.737 ; 0.747
	Burghart et al. (2013)	Goods	101	11	4	58.4%	0.967
	Bruyneel et al. (2014), WP	Goods	42 ; 24 ; 34	9	9	N/A	N/A

Contradicting Experimental Evidence (Giving)

	Trials	*Price Ratios*	*No. of subjects*	*% of GARP satisfiers*	*Average Afriat index*
Fisman, Kariv and Markovits (AER 2007)*	50	Unbounded	76	10.5%	0.108
Andreoni and Miller (ECMT 2002)	8 (8 or 11)	T=3 (T=4)	142 (176)	90.8% (89.8%)	0.003 (0.002)

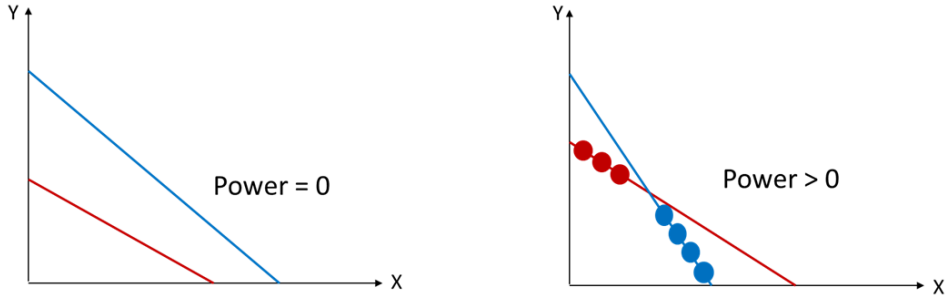
(*) only two-person treatment.





Possible Explanations

- ***Power*** (informally, the probability that random choice fails GARP): affected by the number of intersections between budget lines.



- ***Problem Variability***: affected by the variability in slopes and endowments.
- ***Fatigue***: affected by the number of repetitions and the complexity of the implemented choice rule.
- ***The methodology*** we test textual vs. graphical. Caution: the effect of the methodology on preferences is irrelevant to consistency (is that indeed correct???).

Very Brief Literature Survey

- Most of the literature that is concerned with visual presentation methodologies is focused on risk communication:
  - Some papers consider optimal information presentation (e.g. probabilities in health contexts, managerial data).
  - Other (related) studies show that graphical presentation of lotteries increases risk aversion compared to numerical presentation.
- Harless (1992) claims that some regret effects in the context of binary choice of lotteries are format dependent.
- As far as we know, the literature is restricted to binary choice.

The Setting

- Choice from linear budget sets in the context of other regarding preferences.
- In each decision problem the subject encounters a "modified" dictator game with an anonymous other subject.
- Each token that she allocates to herself is multiplied by α points while a token she allocates to the other is worth β points.

Textual Interface

Divide 68 tokens.

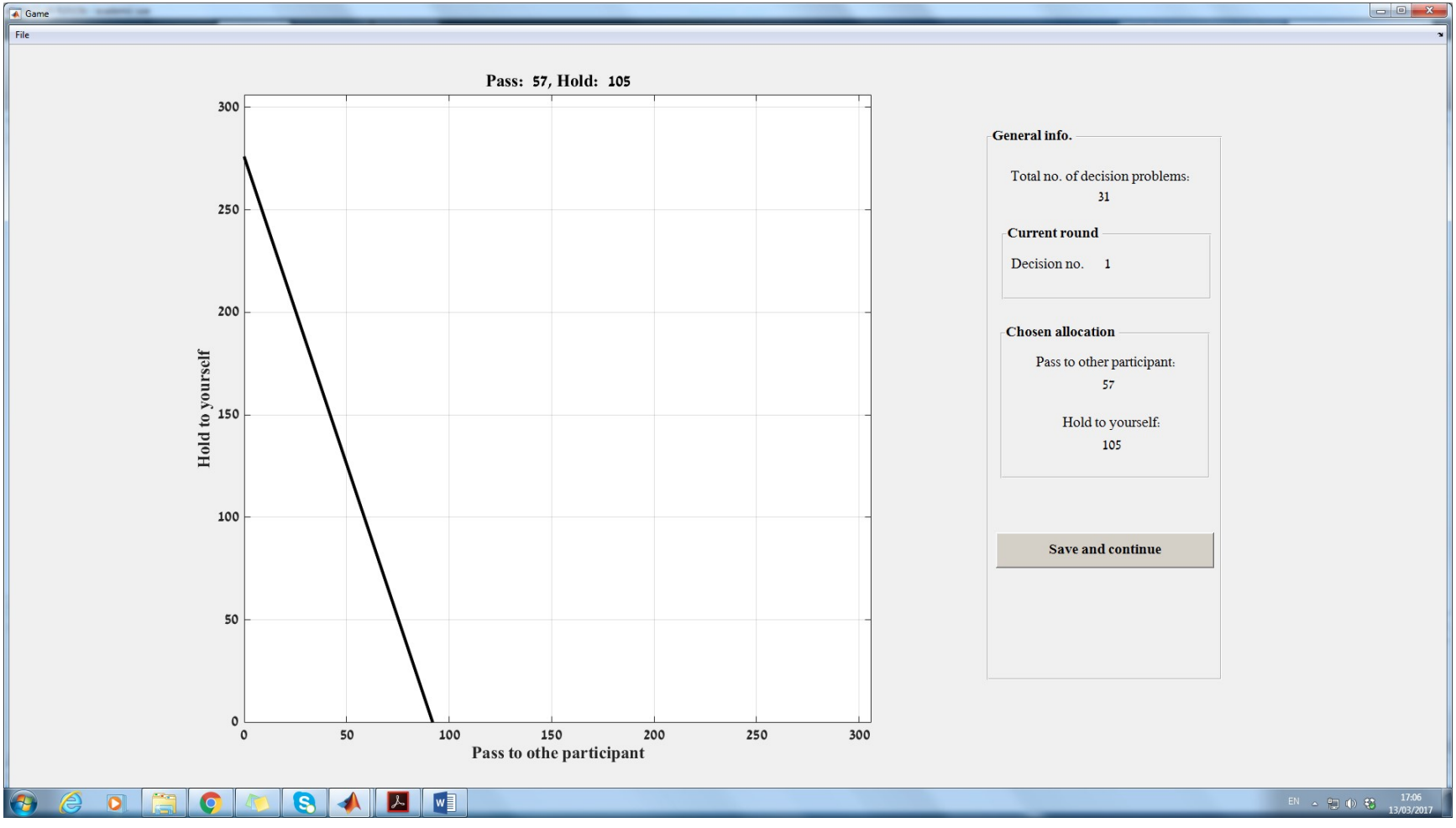
Hold ☐ tokens at 1 point each.

Pass to the other participant _____ tokens at 3 points each.

Decision problem no. 7 of 41

◂ Hebrew ◂ Original

Graphical Interface



◂ Hebrew ◂ Original

Implementation

- At the beginning of the experiment each subject was randomly assigned with:
  - A number of repetitions (between 10 and 50).
  - An upper bound on the price ratio T (between 3 and 12).
- In each trial the subject was randomly assigned with:
  - Price ratio (between $\frac{1}{T}$ and T).
  - Tokens endowment (between 40 and 100).
- Each session was implemented either using the textual methodology (following Andreoni and Miller (2002)) or the graphical methodology (following Fisman et al. (2007) for $n = 2$).
- Monotonicity was imposed in both methodologies.
- Pairs were randomly matched before the experiment, but not revealed to the subjects.

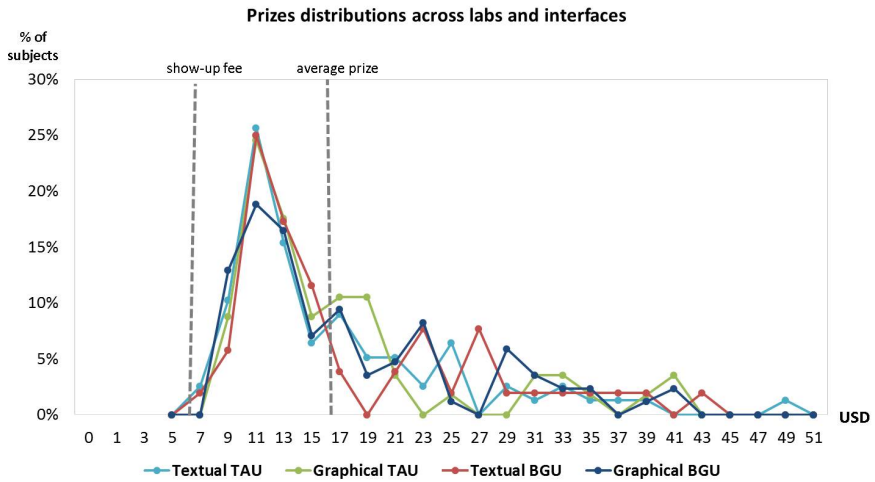
Conversion to Prizes

- Textual Interface:
  - Endowment is tokens.
  - Tokens are converted to points after the DM had made her choice.
  - Points are converted to NIS at the end of the experiment.
- Graphical Interface:
  - The DM chooses a bundle of tokens to hold and to pass.
  - Tokens are converted to NIS at the end of the experiment.
- The conversion rate to NIS was decreasing in T to keep the average prize comparable across treatments.
- The subject's conversion rate was revealed at the beginning of the experiment.
- Participation fees: 25NIS ($\approx 7 USD$).

Subjects and Rewards

- The subjects are 272 undergrads from TAU and BGU.
- The experiments took place between mid March and the end of May, 2016.

	Graphical interface	*Textual interface*
TAU	52	78
BGU	85	57
Total	137	135



Reconstruction

	Trials	*Price Ratios*	*No. of subjects*	*% of GARP satisfiers*	*Average Afriat index*
Fisman, Kariv and Markovits (AER 2007)*	50	Unbounded	76	10.5%	0.108
Andreoni and Miller (ECMT 2002)	8 (8 or 11)	T=3 (T=4)	142 (176)	90.8% (89.8%)	0.003 (0.002)
Graphic interface	41-50	T>8	8	12.5%	0.067
Textual interface	10-29	T=3	10	90%	0

(*) only two-person treatment.

Revealed Preference Relations with Adjustments

The DM chooses bundles $x^i \in \Re^K_+$ $(i \in 1, \ldots, n)$ from budget sets $\{x : p^i x \leq p^i x^i,\ p^i \in \Re^K_{++}\}$.

Let $D = \left\{ \left(p^i, x^i\right)_{i=1}^n \right\}$ be a finite data set, where x^i is the chosen bundle at prices p^i.

Definition

Let $\mathbf{v} \in [0,1]^n$. An observed bundle x^i is

1. **v - Directly Revealed Preferred** to a bundle x, denoted $x^i R^0_{D,\mathbf{v}} x$ if $v^i p^i x^i \geq p^i x$.
2. **v - Strictly Directly Revealed Preferred** to a bundle x, denoted $x^i P^0_{D,\mathbf{v}} x$ if $v^i p^i x^i > p^i x$.
3. **v - Revealed Preferred** to a bundle x, denoted $x^i R_{D,\mathbf{v}} x$ if there exists a sequence of observed bundles $\left(x^j, x^k, \ldots, x^m\right)$ such that $x^i R^0_{D,\mathbf{v}} x^j, x^j R^0_{D,\mathbf{v}} x^k, \ldots, x^m R^0_{D,\mathbf{v}} x$.

Consistency and Rationalizability

Definition

Let $\mathbf{v} \in [0,1]^n$. D satisfies $GARP_{\mathbf{v}}$ if $x^i R_{D,\mathbf{v}} x^j$ implies not $x^j P^0_{D,\mathbf{v}} x^i$.

Definition

Let $\mathbf{v} \in [0,1]^n$. A utility function $u(x)$ $\mathbf{v}-$*rationalizes* D, if for every observed bundle $x^i \in \Re^K_+$, $x^i R^0_{D,\mathbf{v}} x$ implies that $u(x^i) \geq u(x)$.

Afriat's Theorem (1967)

Theorem

The following conditions are equivalent:

1. *There exists a non-satiated utility function that* $\mathbf{1}$*-rationalizes the data.*
2. *The data satisfies* $GARP_{\mathbf{1}}$*.*
3. *There exists a non-satiated, continuous, concave, monotonic utility function that* $\mathbf{1}$*-rationalizes the data.*

Varian Inconsistency Index

Definition

$f_n : [0,1]^n \to [0,M]$, where M is finite, is an *Aggregator Function* if $f_n(\mathbf{1}) = 0$, $f_n(\mathbf{0}) = M$ and $f_n(\cdot)$ is continuous and weakly decreasing.

Definition (Varian Inconsistency Index)

Let $f : [0,1]^n \to [0,M]$ be an aggregator function. *Varian's Inconsistency Index* is,

$$I_V(D,f) = \inf_{\mathbf{v} \in [0,1]^n : D \text{ satisfies } GARP_{\mathbf{v}}} f(\mathbf{v})$$

Other Inconsistency Indices

Definition (Afriat's Critical Cost Efficiency Index)

Let $\mathcal{I} = \left\{ \mathbf{v} \in [0,1]^n : \mathbf{v} = v\mathbf{1}, \forall\, v \in [0,1] \right\}$.

Afriat's Index is, $I_A(D) = \inf_{\mathbf{v} \in \mathcal{I} : D \text{ satisfies } GARP_{\mathbf{v}}} 1 - v$

Definition (Houtman-Maks Inconsistency Index)

Let $f : [0,1]^n \to [0, M]$ be an aggregator function. *Houtman-Maks Inconsistency Index* is,

$$I_{HM}(D, f) = \inf_{\mathbf{v} \in \{0,1\}^n : D \text{ satisfies } GARP_{\mathbf{v}}} f(\mathbf{v})$$

Power

- Bronars (1987):
  - Power is the probability that a DM that chooses randomly (uniformly) on the budget line will fail GARP.
  - Bronars (and others) fail to provide a closed form expression for power in the general case.
  - While understudied, the general intuition is that the power is highly correlated with the number of budget line intersections (which are, in turn, related to the number of trials, the range of slopes and the range of endowments).
  - Bronars (1987) suggests to simulate a large number of such DMs and report frequencies of violations and indices.

Consistency is NOT Procedure Invariant

We use two definitions for consistency:

- Narrow: Those subjects that satisfy GARP.
- Broad: Those subjects that satisfy GARP and those with Afriat inconsistency index equal to epsilon. ◂ Example

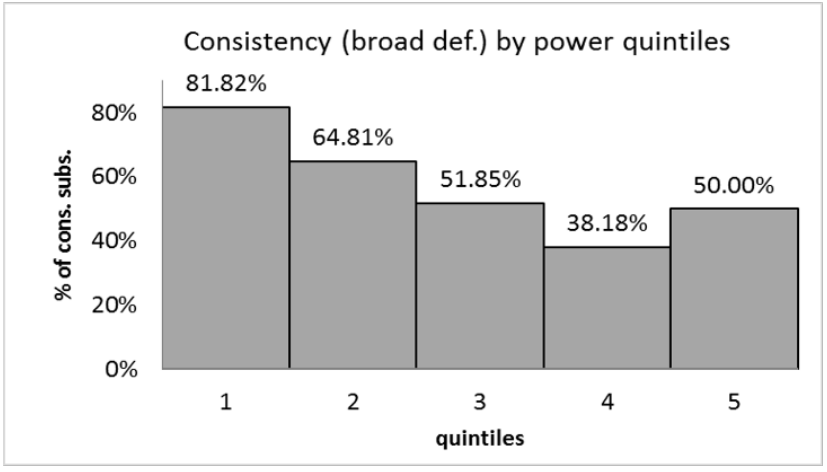
	Narrow definition		Broad definiton		Afriat index	Houtman Maks index
	num. of subs.	% of subs.	num. of subs.	% of subs.		
Textual interface	62	45.9%	69	51.1%	0.11 (0.204)	0.051 (0.094)
Graphical interface	79	57.7%	87	63.5%	0.027 (0.063)	0.029 (0.069)
Total	141	51.8%	156	57.4%	0.068 (0.141)	0.04 (0.083)

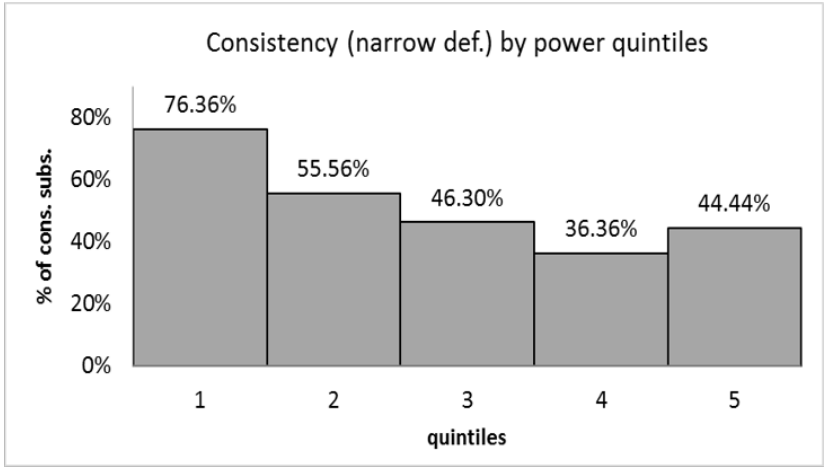
Measuring Power

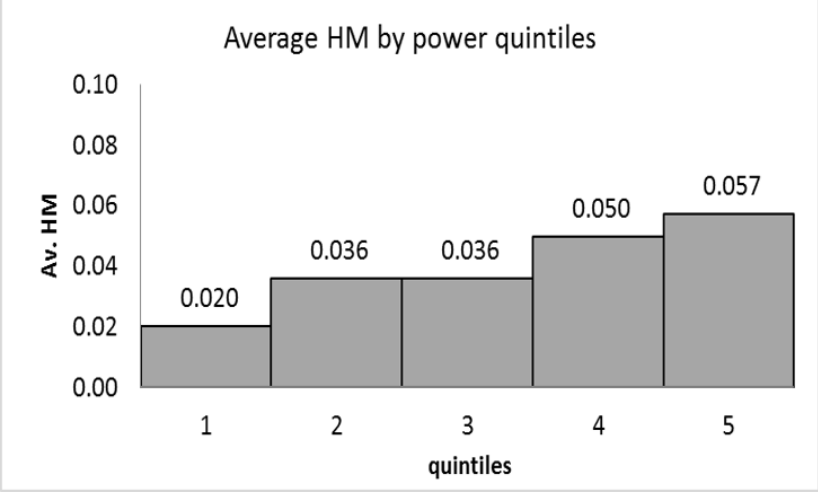
For each subject we ran 10,000 simulations according to Bronars (1987).

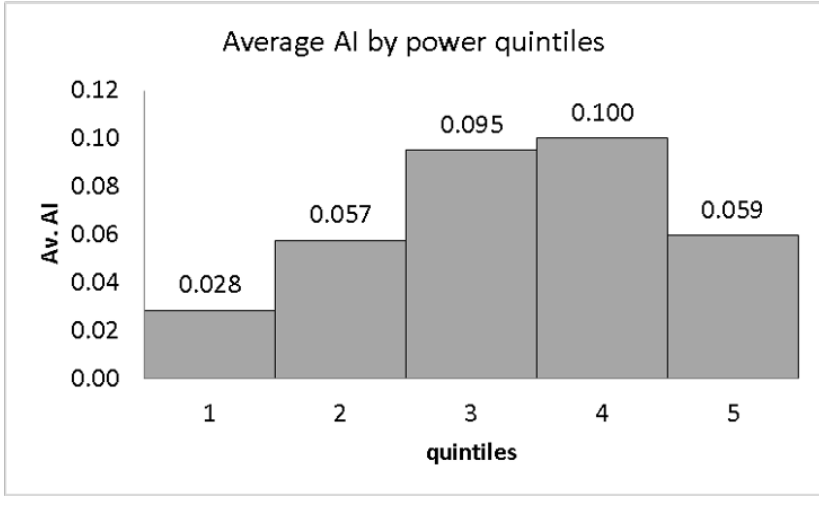
- For each simulation we recorded consistency, number of violations and Afriat inconsistency index.
- We use the median of the number of GARP violations (as percentage of the maximal number).

Does Power Affect Consistency?

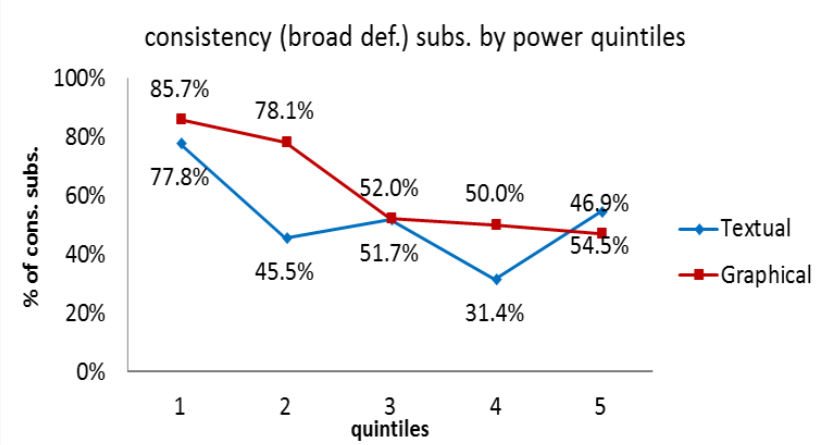


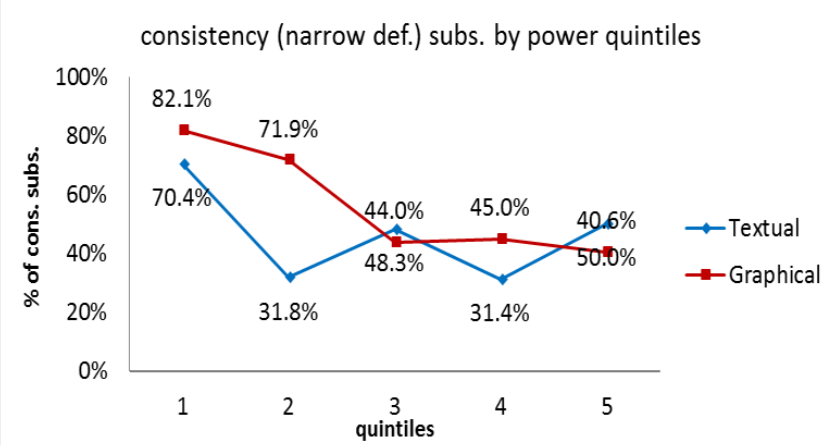


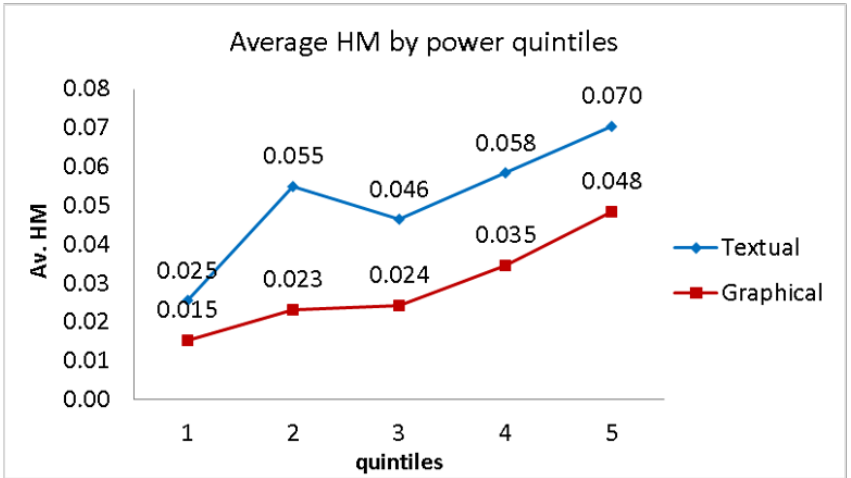


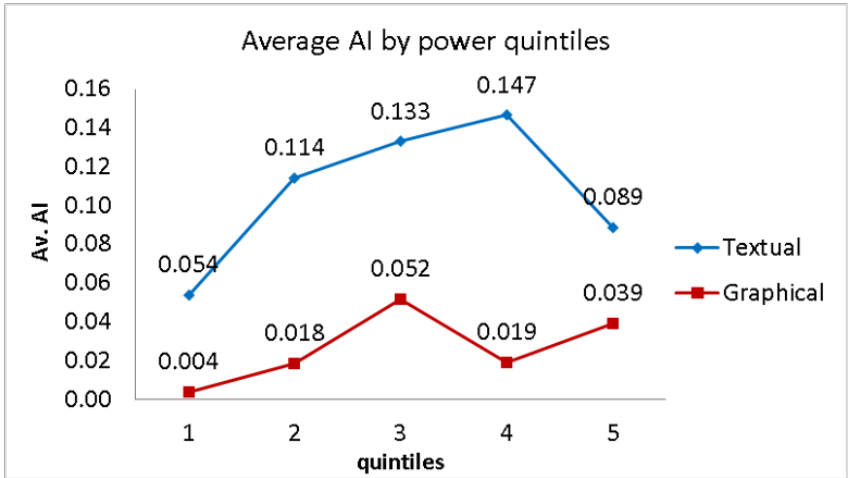


Power Affects Consistency in both Interfaces









Time Measures

- We measured the reaction time for each trial.
- Fatigue: Total RT - the time measured from the beginning of the first trial upto the completion of the last trial (correlation of 0.276 with the number of trials).
- Subjective Complexity: Mean (Median) RT - the Mean (Median) time measured per trial.

The Effect of Total Time

Consistency (broad def.) by time and power

(num. of subs.)

power_quint_5	100.0%	44.4%	57.1%	35.3%	38.5%
	8	9	7	17	13
power_quint_4	75.0%	80.0%	11.1%	11.1%	26.3%
	8	10	9	9	19
power_quint_3	63.6%	60.0%	36.4%	60.0%	41.7%
	11	10	11	10	12
power_quint_2	71.4%	69.2%	68.4%	70.0%	20.0%
	7	13	19	10	5
power_quint_1	95.2%	75.0%	75.0%	77.8%	60.0%
	21	12	8	9	5
	time_quint_1	**time_quint_2**	**time_quint_3**	**time_quint_4**	**time_quint_5**

Consistency (broad def.) by time and power thirds

(Textual, Graphical)

power_third_3	63.6%, 80.0%	53.8%, 50.0%	23.8%, 31.8%
power_third_2	77.8%, 66.7%	33.3%, 55.0%	14.3%, 56.3%
power_third_1	72.7%, 95.2%	63.6%, 80.0%	60.0%, 57.1%
	time_third_1	**time_third_2**	**time_third_3**

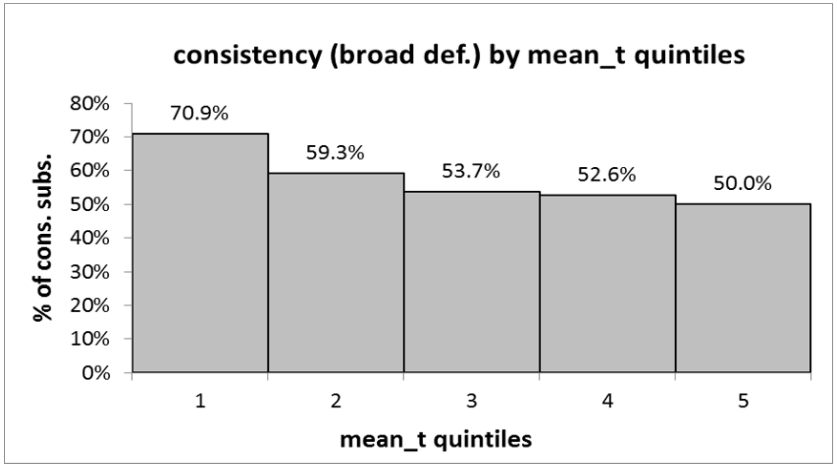
The Effect of Time per Trial

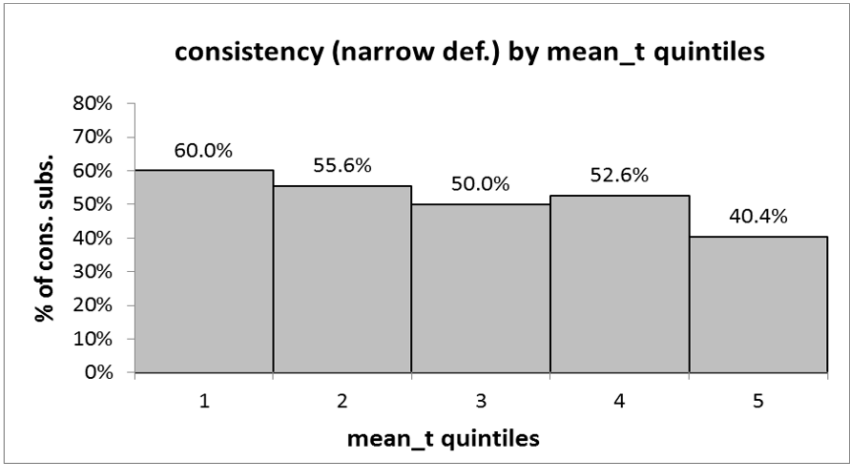
power_quint_5	75.0%	46.2%	23.1%	62.5%	25.0%
	16	13	13	8	4
power_quint_4	58.8%	55.6%	50.0%	18.2%	8.3%
	17	9	6	11	12
power_quint_3	66.7%	52.9%	42.9%	20.0%	53.8%
	12	17	7	5	13
power_quint_2	75.0%	55.6%	64.3%	72.2%	55.6%
	4	9	14	18	9
power_quint_1	83.3%	100.0%	71.4%	84.6%	81.3%
	6	6	14	13	16
	med_rt_quint_1	**med_rt_quint_2**	**med_rt_quint_3**	**med_rt_quint_4**	**med_rt_quint_5**

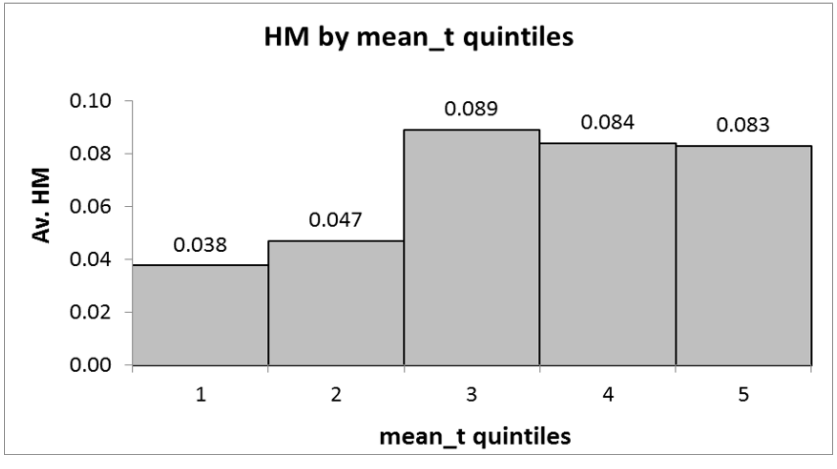
Measuring Problem Variability

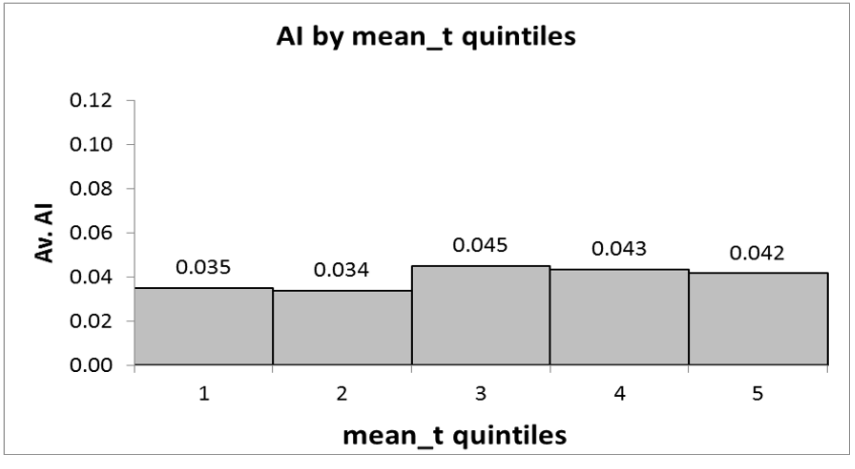
- In the experiment, a maximal slope was randomly assigned to each subject.
- This implies heterogeneity in the variability of the slopes the subjects encounter.
- We measure the problem variability per subject by the mean of the slopes the subject encounters (highly correlated, 0.92, with the standard deviation, by design).

Does Problem Variability Affect Consistency?

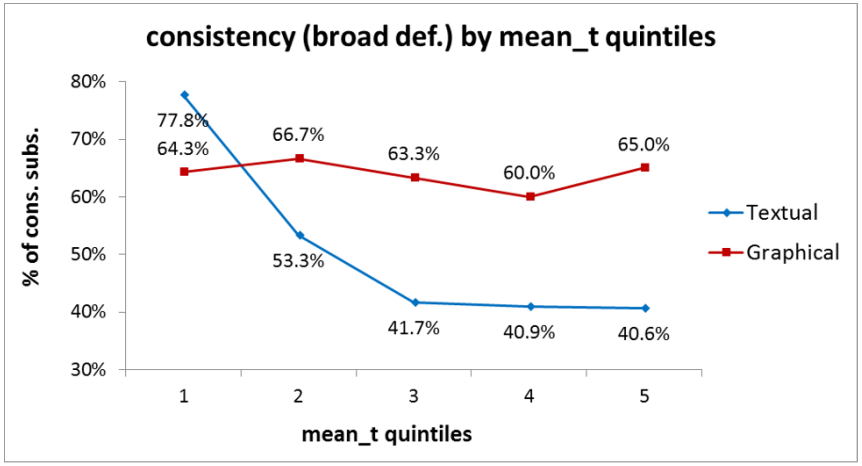


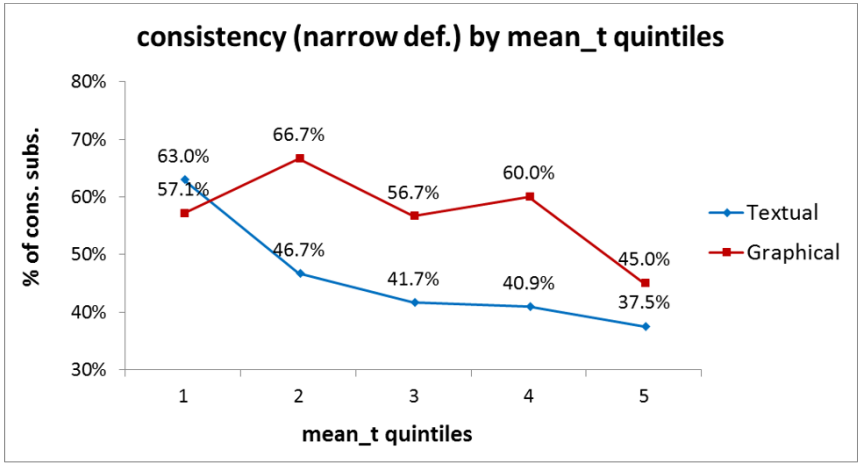


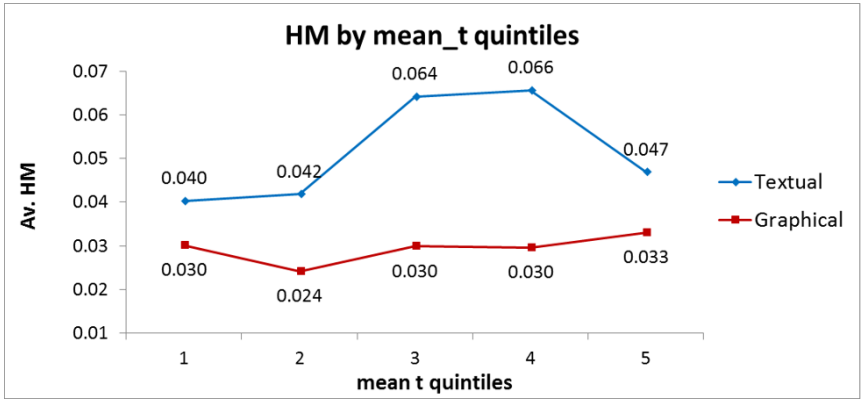


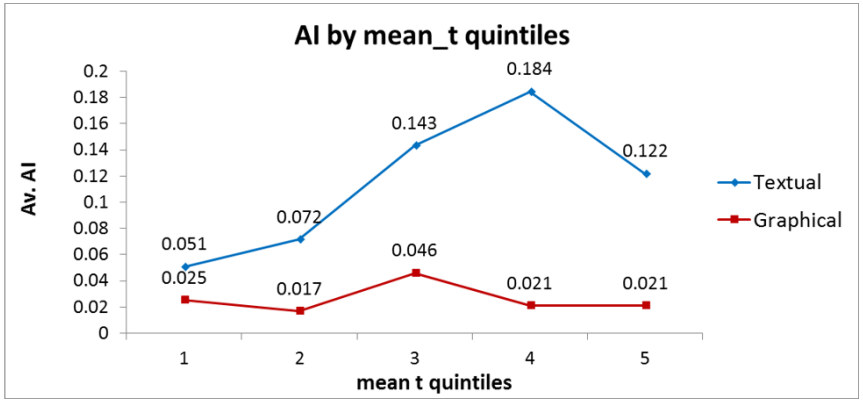


The Effect of Problem Variability by Interface









Consistency Regressions

◂ Narrow Definition

PROBIT

VARIABLES	(1) cons. broad[2] only textual	(2) cons. broad[2] only graphical	(3) cons. broad[2] all subs.	(4) cons. broad[2] all subs.
Interface[2]			**-7.437*** **	**-8.529*** **
			(2.844)	(2.492)
Power	**2.337** **	2.303	2.486	**2.251** **
	(1.086)	(1.624)	(1.583)	(0.890)
*Power * Interface*			-0.264	
			(1.865)	
Average slope	**1.654*** **	0.561	0.515	0.0402
	(0.598)	(0.507)	(0.495)	(0.0921)
*Average slope * Interface*			1.050	**1.519*** **
			(0.743)	(0.566)
Average slope square	**-0.157** **	-0.0608	-0.0547	
	(0.0657)	(0.0575)	(0.0561)	
*Average slope square * Interface*			-0.0936	**-0.147** **
			(0.0827)	(0.0614)
Median RT	-0.0514	-0.00934	-0.00941	-0.0149
	(0.0437)	(0.0443)	(0.0441)	(0.0337)
Median RT Interface*			-0.0368	-0.0301
			(0.0600)	(0.0458)
Time	**0.00442*** **	**0.00321** *	**0.00306** *	**0.00325** **
	(0.00152)	(0.00173)	(0.00169)	(0.00134)
*Time * Interface*			0.000985	0.000746
			(0.00216)	(0.00170)
Gender	-0.438	-0.232	**-0.320** *	**-0.317** *
	(0.286)	(0.263)	(0.191)	(0.191)
Age	**0.174*** **	-0.0451	-0.0422	-0.0461
	(0.0656)	(0.0551)	(0.0541)	(0.0548)
*Age * Interface*			**0.213** **	**0.217*** **
			(0.0829)	(0.0830)
Observations	135	137	272	272
Log Likelihood	-62.536	-70.409	-134.77	-134.278

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0= consistent, 1 = inconsistent

[2] 0= graphical interface, 1 = textual interface

Inconsistency Indices Regressions

TOBIT

VARIABLES	(1) AI only textual	(2) HM only textual	(3) AI only graphical	(4) HM only graphical	(5) AI all subs.	(6) HM all subs.
Interface [1]					**-1.156****	**-0.361****
					(0.457)	(0.147)
Power	0.395	**0.118***	0.287	**0.199***	**0.356****	**0.134****
	(0.249)	(0.0698)	(0.187)	(0.108)	(0.172)	(0.0577)
Average slope	**0.427*****	0.0558	0.0718	-0.00630	0.00626	0.00332
	(0.138)	(0.0367)	(0.0598)	(0.0316)	(0.0196)	(0.00640)
*Average slope * Interface*					**0.359*****	0.0473
					(0.109)	(0.0349)
Average slope square	**-0.0426*****	-0.00556	-0.00752	0.00128		
	(0.0149)	(0.00402)	(0.00678)	(0.00353)		
*Average slope square * Interface*					**-0.0363*****	-0.00506
					(0.0117)	(0.00377)
Median RT	-0.00428	-0.00245	0.00148	0.00200	-0.00110	0.000513
	(0.00927)	(0.00258)	(0.00505)	(0.00296)	(0.00693)	(0.00234)
*Median RT * Interface*					-0.00188	-0.00251
					(0.00883)	(0.00298)
Time	**0.000444***	**0.000209*****	0.000194	0.000158	0.000414	**0.000212****
	(0.000261)	(7.32e-05)	(0.000187)	(0.000111)	(0.000257)	(8.80e-05)
*Time * Interface*					-8.81e-05	-1.52e-06
					(0.000287)	(9.81e-05)
Gender	-0.0306	-0.00712	**-0.0522***	-0.0143	-0.0405	-0.0138
	(0.0616)	(0.0179)	(0.0300)	(0.0174)	(0.0368)	(0.0126)
Age	0.0145	**0.00842****	-0.00787	-0.00393	-0.0113	-0.00470
	(0.0126)	(0.00372)	(0.00636)	(0.00382)	(0.0119)	(0.00373)
*Age * Interface*					0.0207	**0.0122****
					(0.0154)	(0.00505)
Observations	135	135	137	137	272	272
Log Likelihood	-46.384	35.633	-2.649	28.282	-66.261	61.591

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0= graphical interface, 1 = textual interface

The Interface Effect

- The interface has an adverse effect on consistency. Consider a subject of age 24.6 years (average in the sample):
  - For the average slope of 4.2, moving from the graphical interface to the textual interface increases the probability of being inconsistent by 27.1%.
  - For average slopes smaller than 2.42 and larger than 7.415, moving from the graphical interface to the textual interface reduces the probability of being inconsistent.

Other Effects

Holding everything else equal:

- Power (proxy to objective complexity) has an adverse effect on consistency.
- Time spent on the experiment (proxy to fatigue) is negatively correlated with consistency.
- Time spent per trial (proxy to subjective complexity) is not correlated with consistency.
- Accountants are highly consistent ...

Does the Interface Affect Preferences or Heuristics?

- Consistency analysis cannot reveal changes in the distribution of behavior.
- Such analysis requires exploring actual choices rather than their internal consistency.
- We focus on focal types - selfish, altruist, welfare maximizer and egalitarian.
- In addition, we looked into two heuristics based on rounding.
- We naively classify the subjects into these types (or to other).

Focal Types

Types	Andreoni and Miller (ECMT 2002) [1]	Fisman, Kariv and Markovits (AER 2007) [2]	Our experiment			
			Textual		Graphical	
			100% of trials, up to 10% deviation from pure behavior [3]	90% of trials, up to 10% deviation from pure behavior [3]	100% of trials, up to 10% deviation from pure behavior [3]	90% of trials, up to 10% deviation from pure behavior [3]
Selfish	22.7%	26.3%	19.3%	27.4%	27.0%	32.1%
Max. social welfare	6.2%	2.6%	1.5%	5.2%	0.7%	3.6%
Egalitarian	14.2%	2.6%	3.0%	6.0%	0%	1.5%
Altruistic	0%	0%	0%	0%	0%	0%
Round number to self [7]	N/A	N/A	1.5%	7.4% [4,5]	0%	0.7% [4]
Round number to other	N/A	N/A	0%	0%	0% [6]	0% [6]
Other	58.9%	69.5%	74.7%	54%	72.3%	62.1%
Total no. of subjects	**176**	**76**	**135**	**135**	**137**	**137**

[1] *both 8 and 11 trials session.*
[2] *two person treatment only.*
[3] *in the rounding class.., deviation was calculated as 1 token from pure behavior*
[4] *another subject was classified as exhibiting both round to self behavior and max. social welfare behavior*
[5] *about 2/3 of these subjects gave the other subject the remainder from the nearest round number, and another 1/3 gave them larger amounts*
[6] *one subject always gave 1 token to the other subject, and hence was also classified as selfish.*
[7] *54.5% of these subjects were consistent (broad def.)*

Focal Types - Parametric Classification

- In addition, we recover the parameters of a CES utility function for each subject:

$$u(x, y) = [\alpha \times x^{\rho} + (1 - \alpha) \times y^{\rho}]^{\frac{1}{\rho}}$$

  where $\alpha \in [0, 1]$.
- Extreme Altruism: $\alpha = 0$.
- Extreme Selfishness: $\alpha = 1$.
- Egalitarian: $\alpha \in (0, 1)$ and $\rho \to -\infty$.
- Max Social Welfare: $\alpha = \frac{1}{2}$ and $\rho = 1$.
- We recover by the MMI (Halevy et al. (2017)).

Focal Types - Summary

Types	*Andreoni and Miller (ECMT 2002)*	*Fisman, Kariv and Markovits (AER 2007)*	*Our experiment*			
			Textual		*Graphical*	
			100% of trials, up to 10% deviation from pure behavior	*90% of trials, up to 10% deviation from pure behavior*	*100% of trials, up to 10% deviation from pure behavior*	*90% of trials, up to 10% deviation from pure behavior*
Selfish	22.7%	26.3%	19.3%	27.4%	27.0%	32.1%
Max. social welfare	6.2%	2.6%	1.5%	5.2%	0.7%	3.6%
Egalitarian	14.2%	2.6%	3.0%	6.0%	0%	1.5%
Altruistic	0%	0%	0%	0%	0%	0%
Round number to self	N/A	N/A	1.5%	7.4%	0%	0.7%
Round number to other	N/A	N/A	0%	0%	0%	0%
Other	58.9%	69.5%	74.7%	54%	72.3%	62.1%
Total no. of subjects	**176**	**76**	**135**	**135**	**137**	**137**

Types	**Criteria α**	**Criteria ρ**	**Textual**	**Graphical**
Selfish	$\alpha>0.9$		30.4%	38.7%
Max. social welfare	$0.25<\alpha<0.75$	$0.9<\rho<=1$	3.7%	0.7%
Egalitarian		$\rho<-1$	14.8%	6.6%
Altruistic	$\alpha<0.2$		0%	0%
Other			51.1%	54.0%
Total			135	137

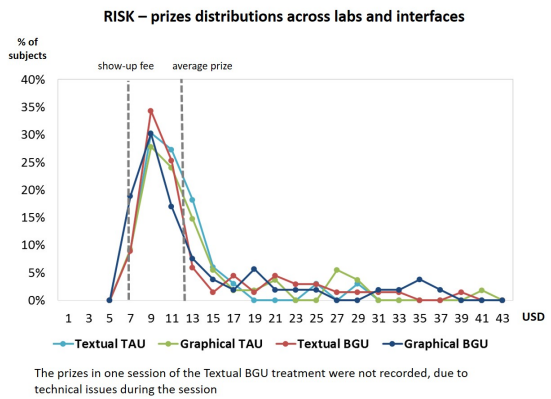
Motivation

- Are the results described so far specific to the modified dictator game settings?
- We compare the two interfaces also in the context of risk.
- Subjects were asked to choose the optimal portfolio of Arrow securities (two equally probable states) from linear budget sets with varying prices (following Choi et al. (2007b)).
- As far as we know, there is no risk preferences experiment using the textual interface.

Subjects and Rewards

- The subjects are 245 undergrads from TAU and BGU.
- The experiments took place between mid November 2016 and the end of January 2017.

	Graphical interface	*Textual interface*
TAU	65	53
BGU	55	72
Total	120	125



The prizes in one session of the Textual BGU treatment were not recorded, due to technical issues during the session

Reconstruction

	Trials	*Price Ratios*	*No. of subjects*	*% of GARP satisfiers*	*Average Afriat index*
Choi, Fisman, Gale and Kariv (AER 2007)*	50	Unbounded	47	25.5%	0.066
Graphic interface	41-50	T>8	13	30.8%	0.04

(*) only symmetric treatment.

Consistency is NOT Procedure Invariant

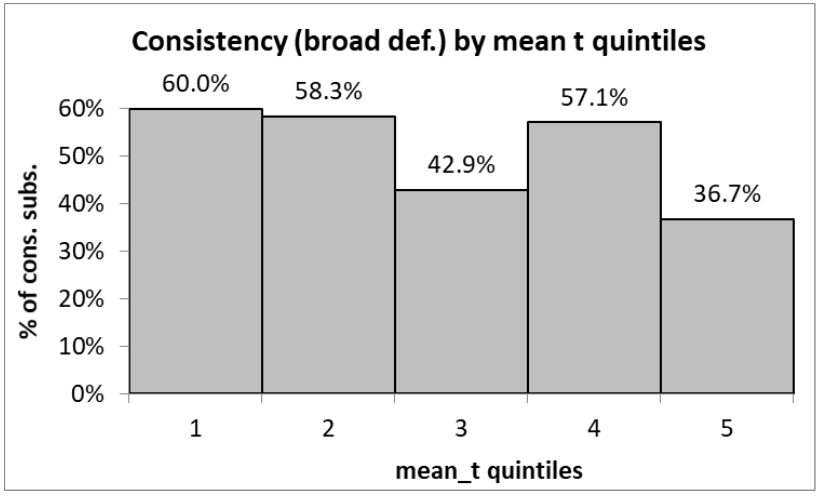
Risk treatment

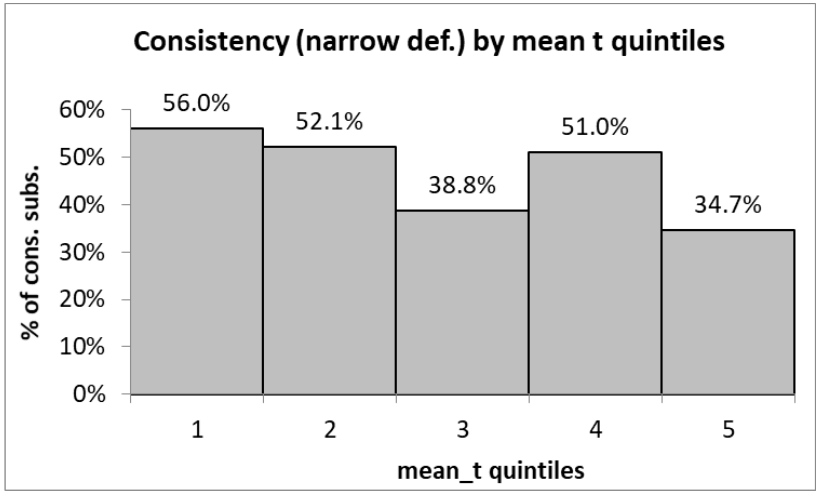
	Narrow definition		Broad definiton		Afriat index	Houtman Maks index
	num. of subs.	% of subs.	num. of subs.	% of subs.		
Textual interface	49	39.2%	58	46.4%	0.044 (0.073)	0.045 (0.074)
Graphical interface	65	54.2%	67	55.8%	0.03 (0.066)	0.032 (0.07)
Total	114	46.5%	125	51.0%	0.037 (0.07)	0.0387 (0.072)

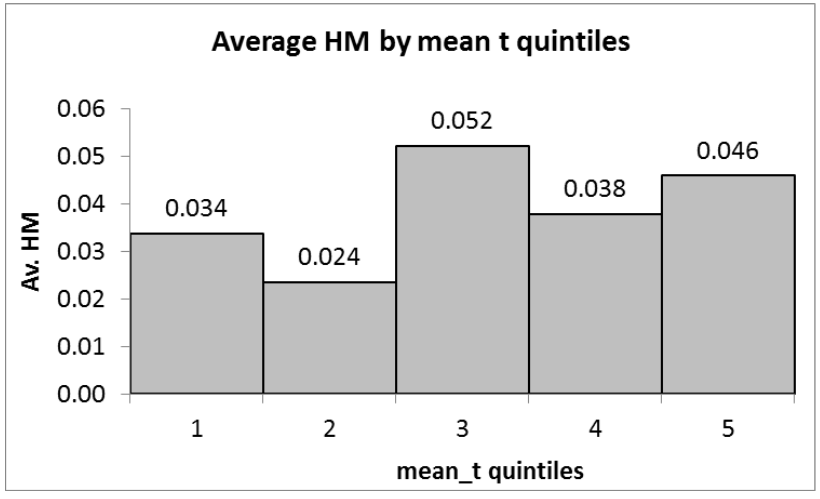
DG treatment

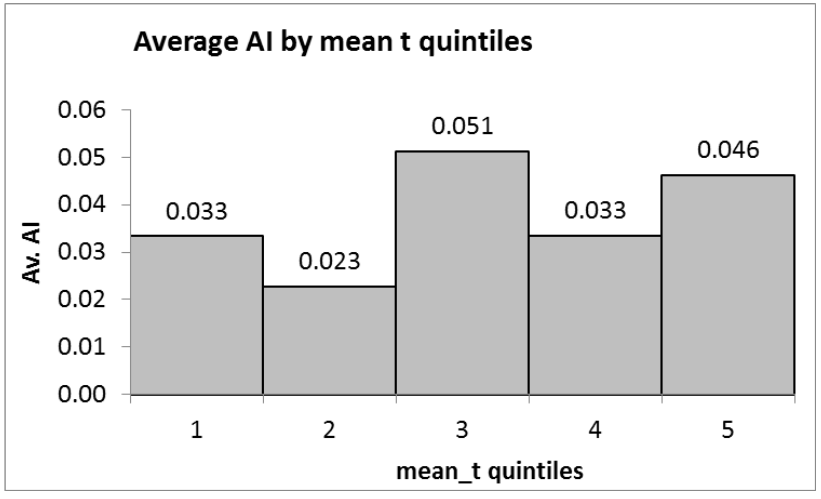
	Narrow definition		Broad definiton		Afriat index	Houtman Maks index
	num. of subs.	% of subs.	num. of subs.	% of subs.		
Textual interface	62	45.9%	69	51.1%	0.11 (0.204)	0.051 (0.094)
Graphical interface	79	57.7%	87	63.5%	0.027 (0.063)	0.029 (0.069)
Total	141	51.8%	156	57.4%	0.068 (0.141)	0.04 (0.083)

Does Problem Variability Affect Consistency?

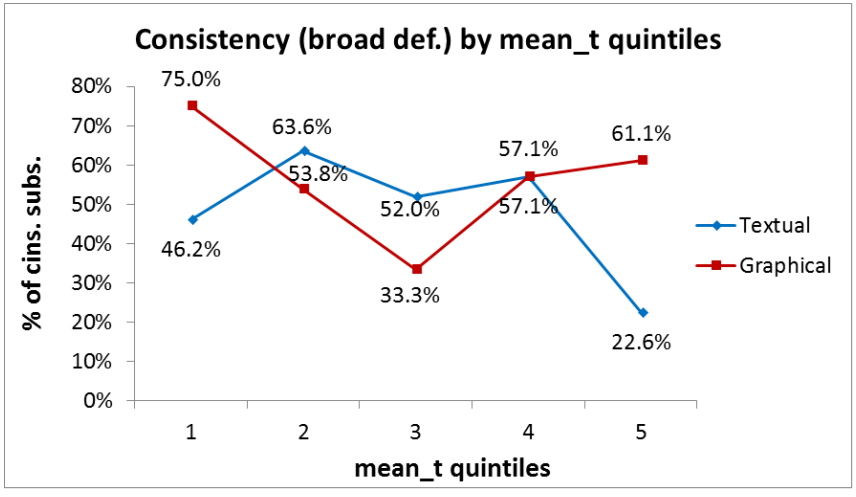


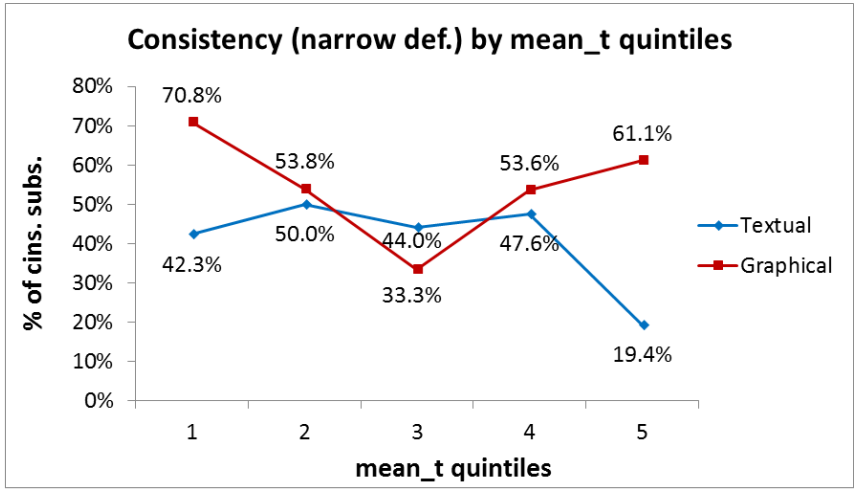


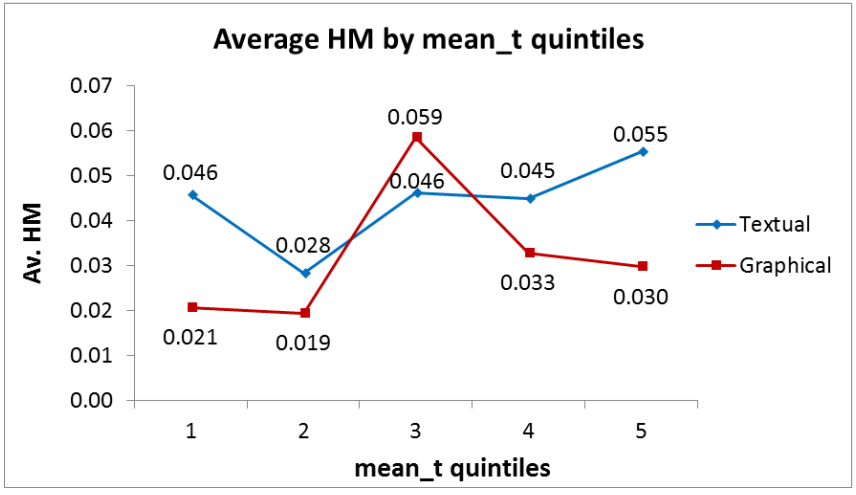


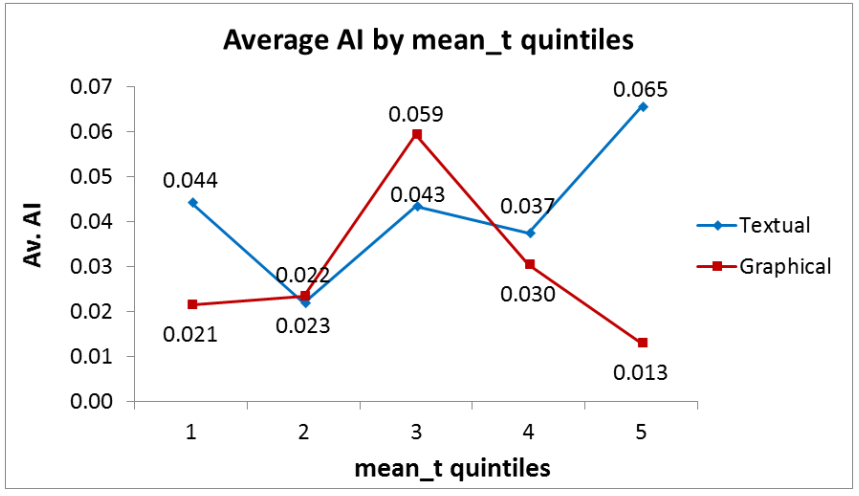


The Effect of Problem Variability by Interface









Consistency Regressions

◂ Narrow Definition

PROBIT

VARIABLES	(1) cons. narrow[1] only textual	(2) cons. narrow[1] only graphical	(3) cons. narrow[1] all subs.
Interface[2]			**4.264***
			(2.420)
Power	**4.022*****	**2.961***	**3.322****
	(1.426)	(1.618)	(1.686)
*Power * Interface*			0.345
			(2.182)
Average slope	**0.291*****	0.0722	**1.718*****
	(0.0979)	(0.106)	(0.647)
*Average slope * Interface*			**-2.361*****
			(0.880)
Average slope square			**-0.200*****
			(0.0765)
*Average slope square * Interface*			**0.306*****
			(0.102)
Median RT	-0.00442	-0.0278	-0.0295
	(0.0349)	(0.0320)	(0.0354)
Median RT			0.0266
			(0.0488)
Time	-0.000279	0.000934	0.000628
	(0.00109)	(0.00101)	(0.00104)
*Time * Interface*			-0.000891
			(0.00148)
Gender	**0.440***	0.0477	0.255
	(0.266)	(0.270)	(0.189)
Age	-0.0115	0.00956	0.00422
	(0.0433)	(0.0377)	(0.0386)
*Age * Interface*			-0.0130
			(0.0577)
Observations	125	118	243
Log Likelihood	-70.055	-63.94	-135.895

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0= consistent, 1 = inconsistent

[2] 0= graphical interface, 1 = textual interface

Inconsistency Indices Regressions

TOBIT

VARIABLES	(1) AI only textual	(2) HM only textual	(3) AI only graphical	(4) HM only graphical	(5) AI all subs.	(6) HM all subs.
Interface [1]					0.198	**0.277***
					(0.213)	(0.146)
Power	**0.255***	**0.155***	**0.304****	**0.302*****	**0.298****	**0.302*****
	(0.130)	(0.0813)	(0.149)	(0.109)	(0.150)	(0.104)
*Power * Interface*					-0.0940	-0.145
					(0.194)	(0.132)
Average slope	**0.0228*****	**0.0112****	**0.189*****	**0.133*****	**0.181*****	**0.126*****
	(0.00855)	(0.00543)	(0.0580)	(0.0419)	(0.0583)	(0.0400)
*Average slope * Interface*					**-0.237*****	**-0.144*****
					(0.0764)	(0.0523)
Average slope square			**-0.0224*****	**-0.0149*****	**-0.0216*****	**-0.0141*****
			(0.00687)	(0.00494)	(0.00694)	(0.00472)
*Average slope square * Interface*					**0.0304*****	**0.0175*****
					(0.00885)	(0.00604)
Median RT	-0.00149	4.78e-05	-0.00101	4.64e-05	-0.00155	2.59e-05
	(0.00311)	(0.00198)	(0.00312)	(0.00216)	(0.00316)	(0.00206)
*median RT * Interface*					-0.000243	-0.000246
					(0.00435)	(0.00289)
Time	5.72e-05	4.58e-05	-6.34e-05	-4.79e-05	-4.44e-05	-4.98e-05
	(9.61e-05)	(6.18e-05)	(9.57e-05)	(6.75e-05)	(9.47e-05)	(6.33e-05)
*Time * Interface*					0.000116	8.32e-05
					(0.000132)	(8.90e-05)
Gender	0.0343	0.0180	0.0184	0.0149	**0.0280***	0.0174
	(0.0232)	(0.0149)	(0.0242)	(0.0176)	(0.0165)	(0.0113)
Age	**0.00709***	0.000368	-0.00122	0.000658	-0.00120	0.000436
	(0.00387)	(0.00247)	(0.00322)	(0.00239)	(0.00327)	(0.00230)
*Age * Interface*					0.00805	0.000354
					(0.00496)	(0.00341)
Observations	125	125	118	118	243	243
Log likelihood	21.649	58.177	15.178	30.099	36.066	86.656

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0 = graphical interface, 1 = textual interface

Focal Types

Risk treatment

Types	Textual		Graphical	
	100% of trials, up to 10% deviation from pure behavior	*90% of trials, up to 10% deviation from pure behavior*	*100% of trials, up to 10% deviation from pure behavior*	*90% of trials, up to 10% deviation from pure behavior*
Cheap corners	6.4%	11.2%	14.2%	21.7%
Safe bundle	12.8%	22.4%	3.3%	5%
Equal shares	0.8%	4%	0%	2.5%
Cutoff	16.8%	20.8%	14.2%	17.5%
Other	63.2%	41.6%	68.3%	53.3%
Total no. of subjects	**125**	**125**	**120**	**120**

DG treatment

Types	*Andreoni and Miller (ECMT 2002)*	*Fisman, Kariv and Markovits (AER 2007)*	*Our experiment*			
			Textual		Graphical	
			100% of trials, up to 10% deviation from pure behavior	*90% of trials, up to 10% deviation from pure behavior*	*100% of trials, up to 10% deviation from pure behavior*	*90% of trials, up to 10% deviation from pure behavior*
Selfish	22.7%	26.3%	19.3%	27.4%	27.0%	32.1%
Max. social welfare	6.2%	2.6%	1.5%	5.2%	0.7%	3.6%
Egalitarian	14.2%	2.6%	3.0%	6.0%	0%	1.5%
Altruistic	0%	0%	0%	0%	0%	0%
Round number to self	N/A	N/A	1.5%	7.4%	0%	0.7%
Round number to other	N/A	N/A	0%	0%	0%	0%
Other	58.9%	69.5%	74.7%	54%	72.3%	62.1%
Total no. of subjects	**176**	**76**	**135**	**135**	**137**	**137**

Focal Types - Parametric Classification

- In addition, we recover the parameters of a DA-CRRA utility function for each subject:

$$u(x,y) = \gamma w(\max\{x,y\}) + (1-\gamma)\, w(\min\{x,y\})$$

where $\gamma = \frac{1}{2+\beta}$ $\quad -1 < \beta < \infty$ and

$$w(x) = \begin{cases} \frac{x^{1-\rho}}{1-\rho} & \rho \geq 0 \quad (\rho \neq 1) \\ ln(x) & \rho = 1 \end{cases}$$

- Corners: either $\beta = -1$ or $\rho = 0$ and $\beta \leq 0$.
- Safe bundle: $\beta \to \infty$.
- Equal shares: $\rho = 1$ and $\beta = 0$
- Cutoff: $\rho = 0$ and $\beta > 0$.
- We recover by the MMI (Halevy et al. (2017)).

Focal Types - Summary

Types	Textual		Graphical	
	100% of trials, up to 10% deviation from pure behavior	*90% of trials, up to 10% deviation from pure behavior*	*100% of trials, up to 10% deviation from pure behavior*	*90% of trials, up to 10% deviation from pure behavior*
Cheap corners	5.6%	11.2%	14.2%	21.7%
Safe bundle	13.6%	23.2%	3.3%	5%
Equal shares	0%	0%	0%	0%
Cutoff	17.6%	20.8%	14.2%	19.2%
Other	63.2%	44.8%	68.3%	54.1%
Total no. of subjects	**125**	**125**	**120**	**120**

Types	Criteria β	Criteria ρ	Textual	Graphical
Cheap corners	$\beta \leq 0$	$\rho \leq 0.2$	6.4%	11.7%
	$\beta < -0.9$	$\rho > 0$		
Safe bundle	$\beta > 2$	$\rho > 0.2$	12.8%	2.5%
Equal shares	$-0.1 < \beta < 0.1$	$0.9 < \rho < 1.1$	1.6%	0%
Cutoff	$\beta > 0$	$\rho \leq 0.2$	28%	15%
Other			51.2%	70.8%
Total			125	120

Summary

- In both contexts: Higher percentage of subjects were consistent when the graphical interface was used (and were less inconsistent).
- In both contexts: The power of the test has an adverse effect on consistency in both interfaces.
- Time spent on the experiment (proxy to fatigue) was negatively correlated with consistency in the dictator game, but not at the risky choice.
- In both contexts: The effect of the slopes differed between interfaces (in a different way).
- The graphical interface seems to encourage corner choices while the textual interface promotes choices on the 45 degree line.

Future

- Two main goals:
  - To improve the classification (High percentage of others).
  - To improve our understanding of the effect of slopes.
- We asked the subjects (post-experiment) to tell us about their decision rule.
- We gave those descriptions to 6 RAs and asked them to:
  - Classify the subjects based on their answers.
  - Classify the subjects based on their choices.
  - Assess the differences.
- We will use these additional data to improve classification and understand the differences between the interfaces (is it indeed harder to implement decision rules in the Textual interface?)
- In addition, we wish to zoom in on the choices of specific subsets of subjects (i.e. those that encountered steep budget lines in the Textual interface in the Dictator game).

Thanks


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
BOLZÏN
MERCI
BÏYAN
SHUKRIA
DANKSCHEEN
TASHAKKUR ATU
SUKSAMA
EKHMET
MEHRBANI
GRAZIE
JUSPAXAR
KOMAPSUMNIDA
GOZAIMASHITA
EFCHARISTO
PALDIES
MAAKE
YAQHANYELAY
TINGKI
SPASSIBO
SNACHALHUYA
NUHUN
CHALTU
NABEEJA
MAITEKA
SPASIBO
DENKAUJA
NENACHALHIYA
SIKOMO
MINMONCHAR
MERASTAWHY
GAEJTHO
TAVTAPUCH
MEDAWAGSE
AGUYJE
FAKAAUE
LAH
ATTO
MAKETAJ

Textual Methodology - Hebrew

back

חלק/י 60 אסימונים.

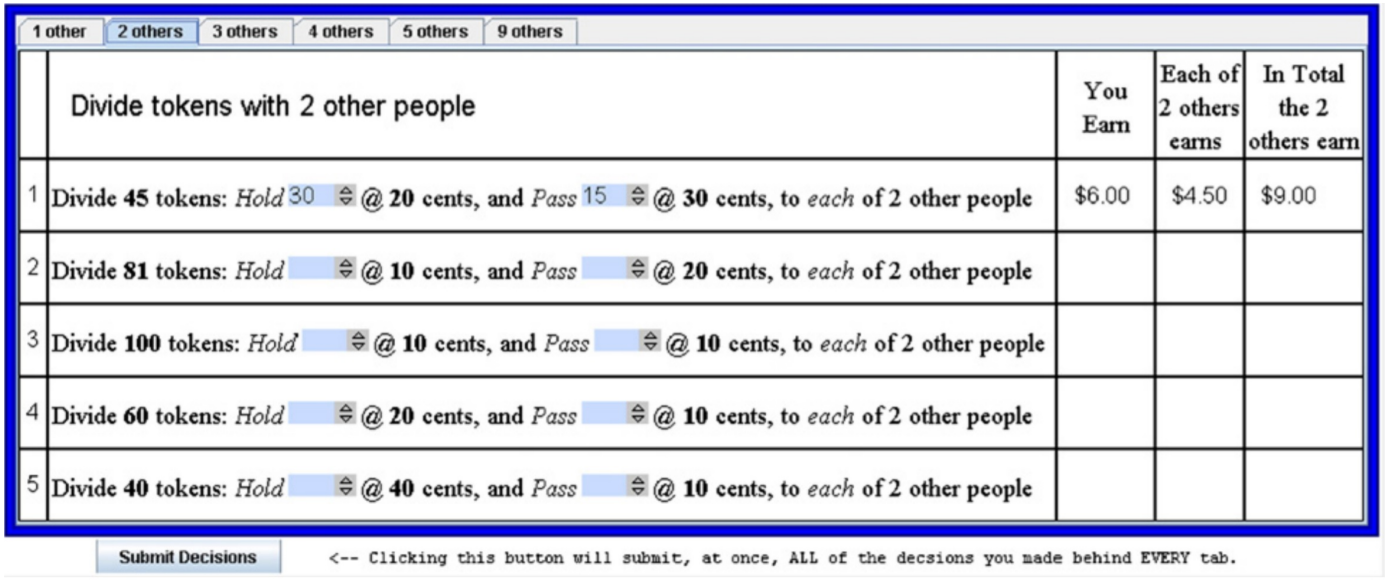
שמור/שמרי לעצמך ☐ אסימונים
בערך של 1 נקודות לאסימון.

העבר/העבירי למשתתף/ת השני/ה _____ אסימונים
בערך של 2 נקודות לאסימון.

בעית החלטה מספר 33
מתוך 34

Textual Methodology - Original

back



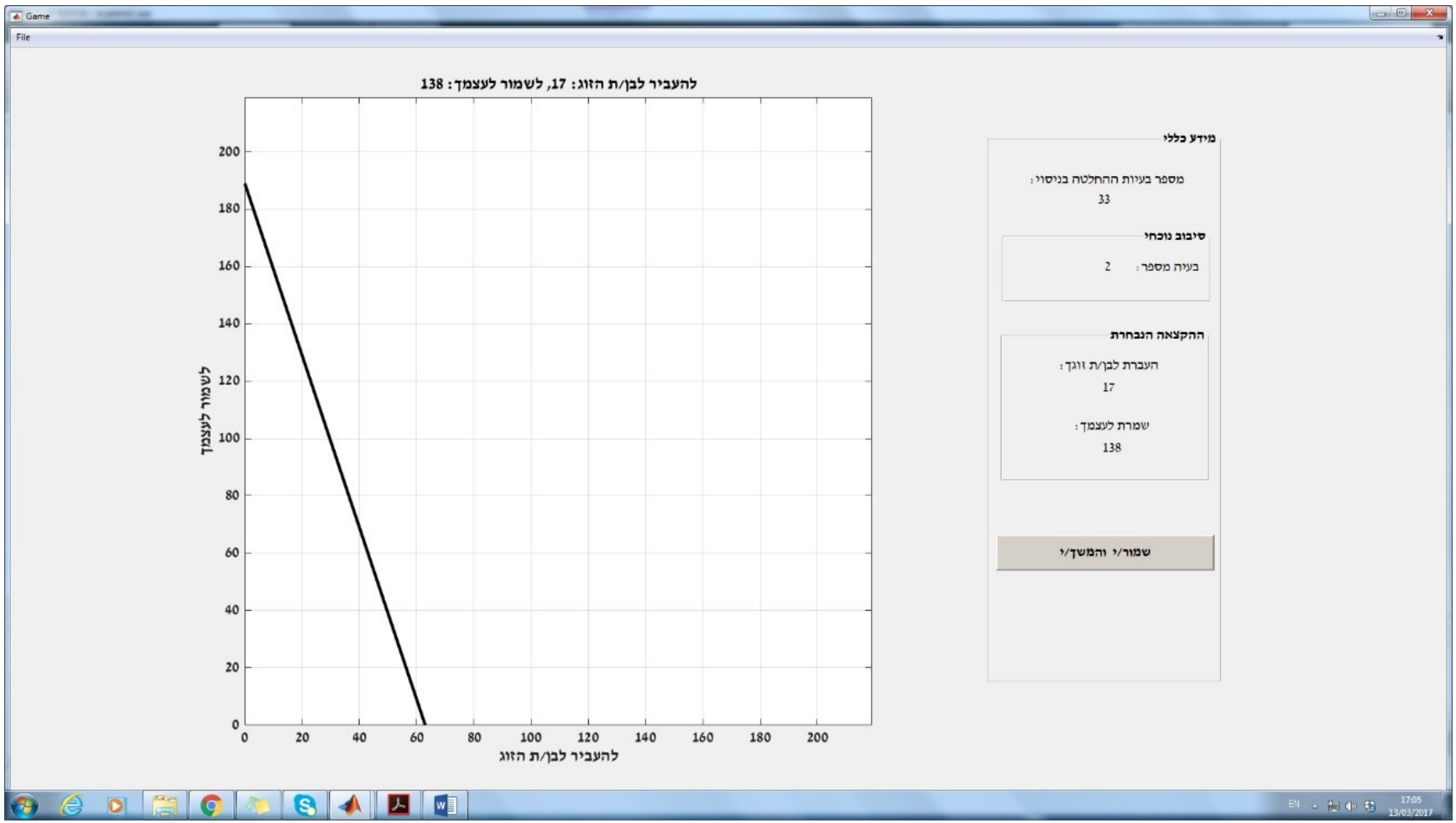
1 other | 2 others | 3 others | 4 others | 5 others | 9 others

	Divide tokens with 2 other people	You Earn	Each of 2 others earns	In Total the 2 others earn
1	Divide **45** tokens: *Hold* 30 @ **20** cents, and *Pass* 15 @ **30** cents, to *each* of 2 other people	$6.00	$4.50	$9.00
2	Divide **81** tokens: *Hold* @ **10** cents, and *Pass* @ **20** cents, to *each* of 2 other people			
3	Divide **100** tokens: *Hold* @ **10** cents, and *Pass* @ **10** cents, to *each* of 2 other people			
4	Divide **60** tokens: *Hold* @ **20** cents, and *Pass* @ **10** cents, to *each* of 2 other people			
5	Divide **40** tokens: *Hold* @ **40** cents, and *Pass* @ **10** cents, to *each* of 2 other people			

Submit Decisions <-- Clicking this button will submit, at once, ALL of the decsions you made behind EVERY tab.

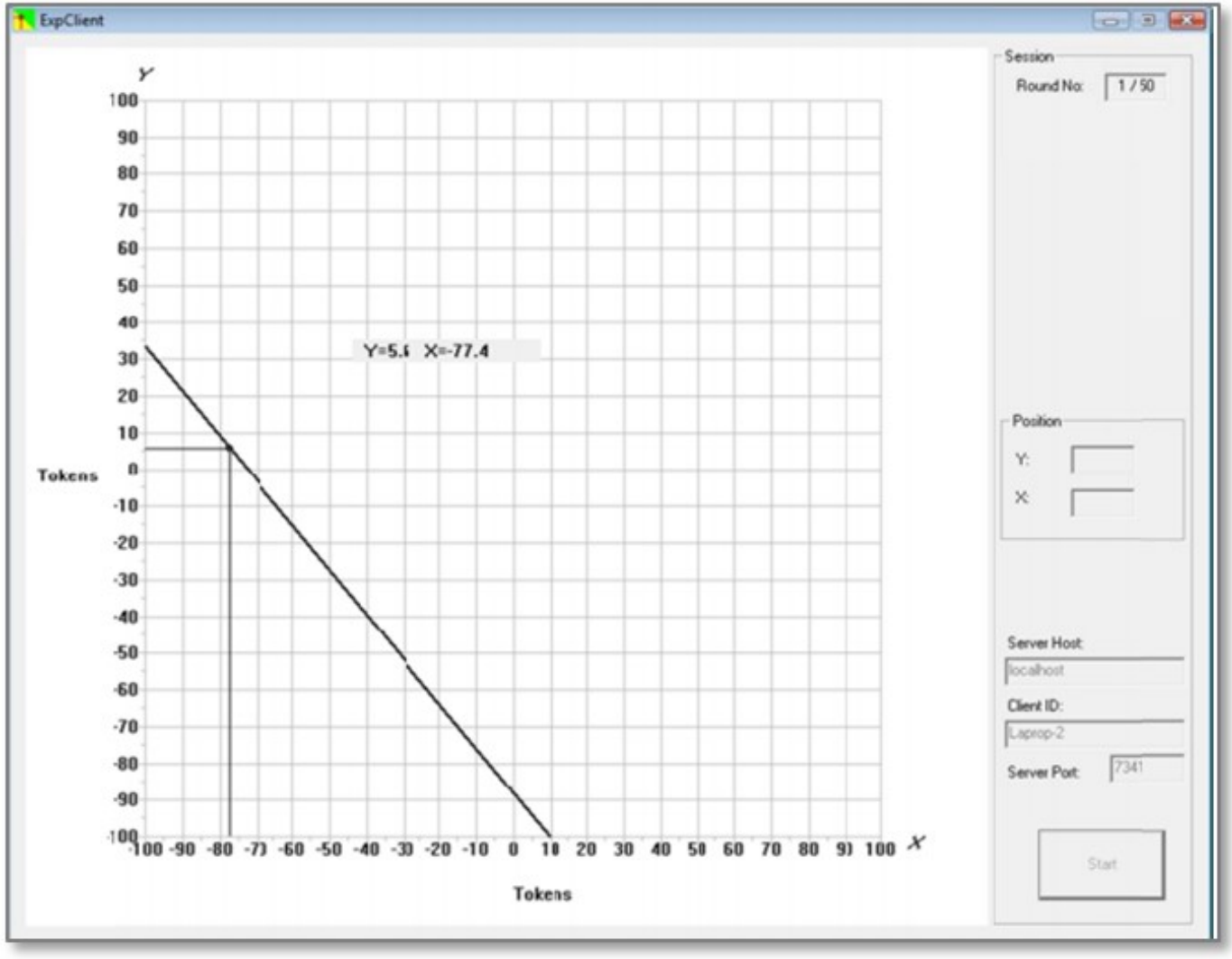
Graphical Methodology - Hebrew

back



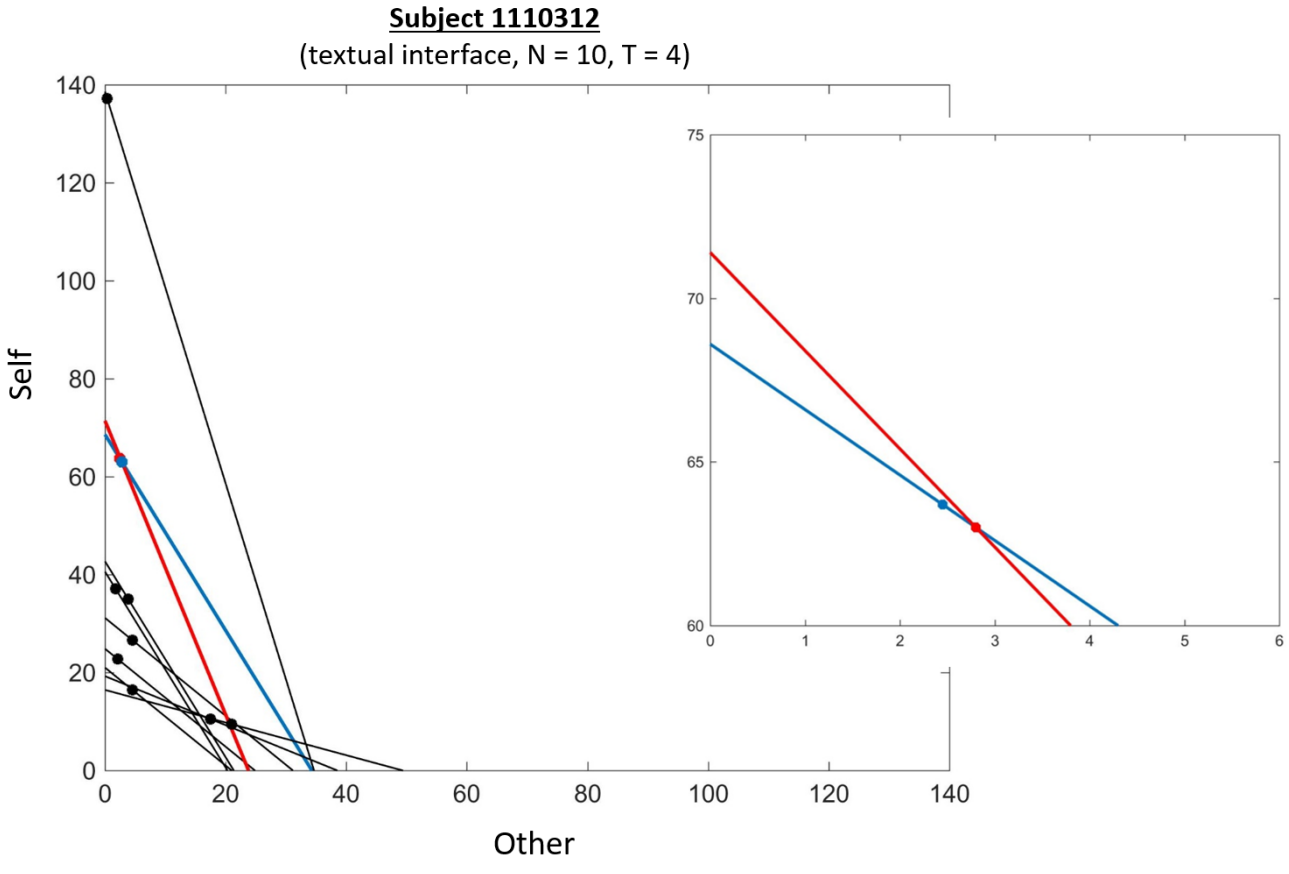
Graphical Methodology - Original

back



Epsilon Afriat Inconsistency Index: Example

back



Consistency Regressions - Narrow Definition

back

PROBIT

VARIABLES	(1) cons. narrow[1] only textual	(2) cons. narrow[1] only graphical	(3) cons. narrow[1] all subs.	(4) cons. narrow[1] all aubs.
Interface[2]			**-6.249****	**-5.735****
			(2.608)	(2.256)
Power	1.538	1.878	2.316	**1.682****
	(1.003)	(1.575)	(1.521)	(0.840)
*Power * Interface*			-0.932	
			(1.781)	
Average slope	**0.842***	-0.270	-0.319	0.0638
	(0.508)	(0.472)	(0.458)	(0.0886)
*Average slope * Interface*			1.042	0.648
			(0.664)	(0.492)
Average slope square	-0.0761	0.0398	0.0472	
	(0.0560)	(0.0535)	(0.0519)	
*Average slope square * Interface*			-0.112	-0.0633
			(0.0743)	(0.0534)
Median RT	-0.0353	-0.0241	-0.0156	-0.0288
	(0.0377)	(0.0439)	(0.0431)	(0.0330)
Median RT			-0.0226	-0.00483
			(0.0562)	(0.0425)
Time	**0.00343*****	**0.00391****	**0.00338****	**0.00384*****
	(0.00130)	(0.00176)	(0.00168)	(0.00134)
*Time * Interface*			5.37e-06	-0.000589
			(0.00205)	(0.00160)
Gender	-0.319	-0.203	-0.270	-0.282
	(0.266)	(0.259)	(0.184)	(0.183)
Age	**0.155****	-0.0486	-0.0405	-0.0369
	(0.0623)	(0.0551)	(0.0536)	(0.0523)
*Age * Interface*			**0.190****	**0.187****
			(0.0796)	(0.0786)
Observations	135	137	272	272
Log Likelihood	-72.569	-72.135	-146.028	-146.618

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0= consistent, 1 = inconsistent

[2] 0= graphical interface, 1 = textual interface

Consistency Regressions - Narrow Definition

back

PROBIT

VARIABLES	(1) Cons. Narrow[1] only textual	(2) Cons. Narrow[1] only graphical	(3) Cons. Narrow[1] all subs.
Interface[2]			3.639
			(2.414)
Power	**3.414****	**3.518****	**3.497****
	(1.371)	(1.691)	(1.667)
*Power * Interface*			-0.00471
			(2.147)
Average slope	**0.233****	**1.726*****	**1.635****
	(0.0971)	(0.652)	(0.639)
*Average slope * Interface*			**-2.131****
			(0.885)
Average slope square		**-0.197*****	**-0.189****
		(0.0766)	(0.0754)
*Average slope square * Interface*			**0.274*****
			(0.103)
Median RT	-0.0118	-0.0172	-0.0205
	(0.0332)	(0.0340)	(0.0338)
*median RT * Interface*			0.0116
			(0.0469)
Time	0.000770	0.000717	0.000584
	(0.00109)	(0.00106)	(0.00102)
*Time * Interface*			-8.47e-05
			(0.00147)
Gender	0.283	-0.0502	0.151
	(0.264)	(0.278)	(0.188)
Age	0.00764	0.00990	0.0109
	(0.0436)	(0.0404)	(0.0394)
*Age * Interface*			0.00287
			(0.0587)
Observations	125	118	243
Log likelihood	-69.003	-65.019	-135.024

Standard errors in parentheses including controls for economics background, lab, academic major, use of calculator (Textual only) and constant

*** p<0.01, ** p<0.05, * p<0.1

[1] 0 = consistent, 1 = inconsistent

[2] 0 = graphical interface, 1 = textual interface