

推論レベルの一貫性: 平均値推測ゲームと金銭要求ゲームによる検証

水 上 英 貴

概 要

本論文では、平均値推測ゲームと金銭要求ゲームをもとに被験者の推論レベルの測定を行い、同一被験者の推論レベルが両ゲーム間で一貫性を持つかどうかを検証した。その結果、両ゲーム間での推論レベルの一貫性は、絶対的推論レベル・相対的推論レベルの両方において、多くの被験者でみられないことがわかった。そこで、どちらのゲームをもとに測定した推論レベルの方が、より一貫性を持つのかも検証した。その結果、金銭要求ゲームをもとに測定した推論レベルの方が、平均値推測ゲームをもとに測定した推論レベルよりも一貫性を持つ可能性が高いことがわかった。

キーワード: Guessing Game, p -Beauty Contest Game, Money Request Game, Level- k Model, 平均値推測ゲーム, p -美人投票ゲーム, 金銭要求ゲーム, レベル k モデル, 推論レベルの一貫性。

JEL Classification Numbers: C72, C91, D03, D83.

1 はじめに

レベル k モデルにしたがって被験者の推論レベル (被験者の合理性) を測定しようと試みるゲームはいくつか知られているが、その中でも有名なゲームは Nagel (1995) による平均値推測ゲーム (Guessing Game) と Arad and Rubinstein (2012) による金銭要求ゲーム (Money Request Game) であろう。

平均値推測ゲームとは、3 人以上のプレイヤーたちがグループとなって 0 以上 100 以下の実数の中から同時に数字を一つ選び、推測値 (同じグループのプレイヤーたちの選択した数字の平均値の p 倍の値) に最も近い数字を選んでいるプレイヤーが勝利するゲームのことである。推測値を決める p の値は実験に

よって様々であるが、Nagel (1995) にしたがって $p = \frac{2}{3}$ で行われることが多い。この平均値推測ゲームを使って様々な被験者の推論レベルを測定しようとする実験には、Camerer (1997), Bosch-Domènech and Nagel (1997), Thaler (1997), Selten and Nagel (1998), Bühren and Frank (2012) などがある¹⁾。平均値推測ゲームを使って被験者の推論レベルを測定すると、これらの実験では、レベル 1 またはレベル 2 である被験者の割合が多く、Nagel (1995) の実験では、推論レベルの最頻値はレベル 1 であった。

金銭要求ゲームとは、2 人のプレイヤーがペアとなって 11 以上 20 以下の整数の中から同時に数字を一つ選び、ペアの相手よりもちょうど 1 だけ小さい数字を選択している場合は自分の選択した数字プラス 20 を、そうでない場合は自分の選択した数字のみをポイントとして得られるゲームのことである。この金銭要求ゲームを使って被験者の推論レベルを測定しようとする実験には Lindner and Sutter (2013) などがあり、金銭要求ゲームの派生版を使った実験も含めると、Alaoui and Penta (2016), Allred et al. (2016), Goeree et al. (2017) などもある。金銭要求ゲームを使って被験者の推論レベルを測定すると、これらの実験では、レベル 1, レベル 2, または、レベル 3 である被験者の割合が多く、Arad and Rubinstein (2012) の実験では、推論レベルの最頻値はレベル 3 であった。

以上のように、平均値推測ゲームや金銭要求ゲームを使って被験者の推論レベルを測定しようとする実験は数多くなされている。そして、どちらのゲームにもとづいて推論レベルを測定してみても、推論レベルがレベル 1, レベル 2, レベル 3 程度の被験者が多いことがわかってきている。

しかしながら、これらの推論レベルの測定実験は別々の被験者によって行われたものであり、同一の被験者によって行われたものではない。したがって、平均値推測ゲームによる推論レベルと金銭要求ゲームによる推論レベルが同一被験者にとっては同一レベルになるのかという、異なるゲーム間での推論レベルの一貫性についてはまだほとんど検証がされていない。あるゲームで測定し

1) 実験内容や結果の概要は、それぞれの論文、または、Nagel (2008) や水上 (2015) などを参照されたい。

た推論レベルが低いにもかかわらず、別のゲームで測定した推論レベルが高くなってしまえば、測定するゲームに被験者の推論レベルが依存してしまい、測定された推論レベルの信頼性が低くなる。よって、推論レベルが被験者に固有なものであり、ゲームの種類によらず一貫性を持つのかどうかを明らかにすることは、重要な課題の一つであろう。

1.1 関連文献

同一被験者の異なるゲーム間での推論レベルの一貫性について検証した研究には、Dittrich and Leipold (2014), Georganas et al. (2015) などがある。Dittrich and Leipold (2014) は、平均値推測ゲームおよび金銭要求ゲームを使って被験者の推論レベルをそれぞれ測定し、被験者の両ゲーム間での推論レベルの一貫性について検証した。その結果、80.28% の被験者について、両ゲーム間で推論レベルの一貫性がみられることを明らかにした。しかしながら、Dittrich and Leipold (2014) では、しばしば使われる被験者の選択結果ではなく、実験後にそれぞれの被験者に説明をさせた選択理由を使って、推論レベルの測定を行っていた。そのため、平均値推測ゲームでは 88.94%、金銭要求ゲームでは 81.67% もの被験者の推論レベルがレベル 0 となっていた。しかしながら、Nagel (1995) によれば、平均値推測ゲームでは推論レベル 1 が、Arad and Rubinstein (2012) によれば、金銭要求ゲームでは推論レベル 3 が推論レベルの最頻値である。したがって、被験者の選択結果ではなく選択理由を使って推論レベルの測定を行なった Dittrich and Leipold (2014) の測定手法には問題があり、両ゲーム間での推論レベルの一貫性がかなり多くの被験者でみられたという結果は再検証される必要があるだろう。

一方、Georganas et al. (2015) は、Costa-Gomes and Crawford (2006) による二人平均値推測ゲーム (Two-Person Guessing Game) と切り下げゲーム (Undercutting Game) にもとづいて被験者の推論レベルをそれぞれ測定した。なお、二人平均値推測ゲームは平均値推測ゲームの派生版であり、切り下げゲームは Georganas et al. (2015) によるものであるが金銭要求ゲームと似たゲームである。Georganas

et al. (2015) は、ほとんどの被験者について、二人平均値推測ゲームと切り下げゲーム間で、推論レベルの一貫性はみられないことを示した。さらに、被験者の推論レベルに、相対的な一貫性もみられないことも示した²⁾。

その他に、平均値推測ゲームと金銭要求ゲームの両方を同一被験者にプレイさせた研究として、Sousa et al. (2014) がある。Sousa et al. (2014) は、チェス・プレイヤーを被験者として、平均値推測ゲームを使って被験者を戦略的なプレイヤーと非戦略的なプレイヤーに分類し、戦略的なプレイヤーの方が非戦略的なプレイヤーと比べて金銭要求ゲームにおいて混合戦略ナッシュ均衡戦略をプレイしていることを示した。

1.2 論文の目的と構成

本論文の目的は、Dittrich and Leipold (2014), Georganas et al. (2015) と同様に、二つのゲームを使って被験者の推論レベルをそれぞれ測定し、異なるゲーム間での推論レベルの一貫性について検証することである。本論文では、しばしばなされるように被験者の選択結果にもとづいて、平均値推測ゲームおよび金銭要求ゲームにおける被験者の推論レベルを測定し、Dittrich and Leipold (2014) の結果の再検証を行う。

本論文では、Dittrich and Leipold (2014) とは違って、被験者の選択結果を使って推論レベルの測定を行った結果、異なるゲーム間での推論レベルの一貫性がみられたのは、被験者のわずか 22.75% であった。この値の低さの一因は、被験者の選択結果にもとづいて推論レベルの測定を行う際の、平均値推測ゲームと金銭要求ゲームでの推論レベルの分布の違いにある。先に述べたように、平均値推測ゲームでの被験者の推論レベルの最頻値はレベル 1 であり、金銭要求ゲームでの被験者の推論レベルの最頻値はレベル 3 である。それゆえ、両方のゲームで最頻値の推論レベルをとる被験者でさえも、推論レベルの一貫性を持

2) Georganas et al. (2015) は、あるゲームで個人 i の推論レベルが個人 $j \neq i$ よりも高ければ、別のゲームでも個人 i の推論レベルが個人 j よりも高くなることを、推論レベルの相対的な一貫性とよんでいる。

たないことになる。しかしながら、両方のゲームで最頻値の推論レベルをとる被験者は、推論レベルが一貫していると考えられるだろう。

そこで、本論文では、相対的推論レベルに着目し、相対的推論レベルの一貫性についても検証する。その結果、被験者の選択結果を使って推論レベルの測定を行ったところ、異なるゲーム間での相対的推論レベルの一貫性がみられたのは、被験者の 37.73% であった。この値も、Dittrich and Leipold (2014) の 81.67% よりも低くなっている。以上のことから、Dittrich and Leipold (2014) の結果を再検証するために、被験者の選択理由ではなく選択結果を使って被験者の推論レベルの測定を行ったところ、平均値推測ゲームと金銭要求ゲーム間で、絶対的推論レベル・相対的推論レベルの両方において、推論レベルの一貫性はみられないことがわかった。そこで、どちらのゲームをもとに測定した推論レベルの方が、より一貫性を持つのかも検証していく。

本論文の構成は以下のとおりである。まず、平均値推測ゲームと金銭要求ゲームの理論モデルを 2 章で紹介する。そして、本研究で実施した実験内容について 3 章で、その実験結果について 4 章で紹介する。その後、5 章で実験結果の分析を行い、被験者の平均値推測ゲームと金銭要求ゲーム間での推論レベルの一貫性について検証していく。

2 理論モデル

2.1 平均値推測ゲーム

$N^G := \{1, 2, \dots, n\}$ をプレイヤーの集合とし、 $n \geq 3$ とする。プレイヤー $i \in N^G$ の行動を a_i^G 、 n 人のプレイヤーの行動の組を $a^G = (a_1^G, a_2^G, \dots, a_n^G)$ と表現する。

平均値推測ゲームでは、 n 人のプレイヤーがグループとなり、それぞれのプレイヤー $i \in N^G$ は、閉区間 $[0, 100]$ から数字 (実数) を一つ選択する。したがって、平均値推測ゲームでの行動 a_i^G は、それぞれのプレイヤー $i \in N^G$ について、 $a_i^G \in [0, 100]$ となる。 n 人のプレイヤーの行動の組 $a^G = (a_1^G, a_2^G, \dots, a_n^G)$ に対

して、その平均値 m を $m := \frac{1}{n} \sum_{i=1}^n a_i^G$ と定義するとき、 $g := \frac{2}{3}m$ を推測値とよぶ。平均値推測ゲームにおけるプレイヤー $i \in N^G$ の利得 $u_i(a^G)$ は、

$$u_i(a^G) = C - c \left| g - a_i^G \right|$$

である。ただし、 $C (> 0)$ は推測値を正しく予測できた報酬、 $c (> 0)$ は推測値を正しく予測できなかった罰金である。

本論文では、レベル k モデルとして、Nagel (1995) および Duffy and Nagel (1997) による繰り返し最適反応 (Iterated Best Response) モデルを使って、平均値推測ゲームにおけるプレイヤーたちの推論レベルを決定する。レベル k モデルでは、非戦略的なプレイヤーをレベル 0 と定義するが、平均値推測ゲームでは、非戦略的なプレイヤーは閉区間 $[0, 100]$ からランダムに数字を選択すると仮定されることが多い。また、繰り返し最適反応モデルでは、プレイヤーが基準点に対して繰り返して何回最適反応をとっているかでプレイヤーの推論レベル k ($k = 1, 2, \dots, +\infty$) を表現する。本論文でも、レベル 0 のプレイヤーは閉区間 $[0, 100]$ からランダムに数字を選択すると仮定するため、それらの平均値である数字 50 が基準点となる。

本論文では、プレイヤーたちの行動の分析が簡単になることもあり、Nagel (1995) および Duffy and Nagel (1997) にしたがって、推論レベルを区間で表現する³⁾。

定義 1 (平均値推測ゲームにおける推論レベル). 推論レベルを k ($k = 1, 2, \dots, +\infty$) とする。このとき、プレイヤー $i \in N^G$ の行動 a_i^G が

$$a_i^G \in \left[50 \cdot \left(\frac{2}{3} \right)^{k+\frac{1}{2}}, 50 \cdot \left(\frac{2}{3} \right)^{k-\frac{1}{2}} \right)$$

のときに推論レベル k の行動、

$$a_i^G \in \left[50 \cdot \left(\frac{2}{3} \right)^{\frac{1}{2}}, 100 \right]$$

3) 推論レベルを区間で表現することについては、水上 (2015) も参照されたい。

推論レベルの一貫性

のときに推論レベル 0 の行動とよぶ。

以上より、平均値推測ゲームにおける推論レベル別のプレイヤーの行動は、表 1 にまとめられる。また、平均値推測ゲームのナッシュ均衡は定理 1 のとおりである。

レベル	行動
レベル 0	[40.82, 100]
レベル 1	[27.22, 40.82)
レベル 2	[18.14, 27.22)
レベル 3	[12.10, 18.14)
レベル 4	[8.06, 12.10)
レベル 5	[5.38, 8.06)
⋮	⋮
レベル ∞	0

表 1 推論レベル別のプレイヤーの行動 (平均値推測ゲーム)

定理 1. 平均値推測ゲームでは、すべてのプレイヤーが 0 を選択する行動がナッシュ均衡となる。

2.2 金銭要求ゲーム

$N^{\text{MR}} := \{1, 2\}$ をプレイヤーの集合とする。プレイヤー $i \in N^{\text{MR}}$ の行動を a_i^{MR} 、2 人のプレイヤーの行動の組を $a^{\text{MR}} = (a_1^{\text{MR}}, a_2^{\text{MR}})$ と表現する。

金銭要求ゲームでは、2 人のプレイヤーがペアとなり、それぞれのプレイヤー $i \in N^{\text{MR}}$ は、整数の集合 $\{z \in \mathbb{Z} \mid 11 \leq z \leq 20\}$ から数字 (整数) を一つ選択する。したがって、金銭要求ゲームでの行動 a_i^{MR} は、それぞれのプレイヤー $i \in N^{\text{MR}}$ について、 $a_i^{\text{MR}} \in \{z \in \mathbb{Z} \mid 11 \leq z \leq 20\}$ となる。金銭要求ゲームにおけるプレイヤー $i \in N^{\text{MR}}$ の利得 $u_i(a^{\text{MR}})$ は、

$$u_i(a^{\text{MR}}) = \begin{cases} a_i^{\text{MR}} + 20 & a_i^{\text{MR}} = a_j^{\text{MR}} - 1 \text{ のとき} \\ a_i^{\text{MR}} & a_i^{\text{MR}} \neq a_j^{\text{MR}} - 1 \text{ のとき} \end{cases}$$

である。つまり、ベアの相手よりもちょうど 1 だけ小さい整数を選択している場合は選択した整数プラス 20 の利得となり、そうでない場合は選択した整数のみの利得となる。

本論文では、レベル k モデルとして、Arad and Rubinstein (2012) にしたがって、金銭要求ゲームにおけるプレイヤーたちの推論レベルを決定する。Arad and Rubinstein (2012) では、金銭要求ゲームにおける非戦略的行動は 20 を選択することとしているため、本論文でも 20 を選択するプレイヤーをレベル 0 とする。したがって、最適反応の基準点は 20 となる。

定義 2 (金銭要求ゲームにおける推論レベル). 推論レベルを k ($k = 0, 1, \dots, 9$) とする。このとき、プレイヤー $i \in N^{\text{MR}}$ の行動 a_i^{MR} が

$$a_i^{\text{MR}} = 20 - k$$

のときに**推論レベル k の行動**とよぶ。

したがって、金銭要求ゲームにおける推論レベル別のプレイヤーの行動は、表 2 にまとめられる。

レベル	行動
レベル 0	20
レベル 1	19
レベル 2	18
レベル 3	17
レベル 4	16
レベル 5	15
⋮	⋮
レベル 9	11

表 2 推論レベル別のプレイヤーの行動 (金銭要求ゲーム)

さらに、金銭要求ゲームを利得行列で表現すると、表 3 のようになる。したがって、金銭要求ゲームのナッシュ均衡は定理 2 のとおりである。

定理 2. 金銭要求ゲームでは、どちらのプレイヤーも、11 を 0, 12 を 0, 13 を 0, 14 を 0, 15 を 0.25, 16 を 0.25, 17 を 0.2, 18 を 0.15, 19 を 0.1, 20 を 0.05

推論レベルの一貫性

		プレイヤー 2									
		11	12	13	14	15	16	17	18	19	20
プレイヤー 1	11	11, 11	21, 12	11, 13	11, 14	11, 15	11, 16	11, 17	11, 18	11, 19	11, 20
	12	12, 21	12, 12	22, 13	12, 14	12, 15	12, 16	12, 17	12, 18	12, 19	12, 20
	13	13, 11	13, 22	13, 13	23, 14	13, 15	13, 16	13, 17	13, 18	13, 19	13, 20
	14	14, 11	14, 12	14, 23	14, 14	24, 15	14, 16	14, 17	14, 18	14, 19	14, 20
	15	15, 11	15, 12	15, 13	15, 24	15, 15	25, 16	15, 17	15, 18	15, 19	15, 20
	16	16, 11	16, 12	16, 13	16, 14	16, 25	16, 16	26, 17	16, 18	16, 19	16, 20
	17	17, 11	17, 12	17, 13	17, 14	17, 15	17, 26	17, 17	27, 18	17, 19	17, 20
	18	18, 11	18, 12	18, 13	18, 14	18, 15	18, 16	18, 27	18, 18	28, 19	18, 20
	19	19, 11	19, 12	19, 13	19, 14	19, 15	19, 16	19, 17	19, 28	19, 19	29, 20
	20	20, 11	20, 12	20, 13	20, 14	20, 15	20, 16	20, 17	20, 18	20, 29	20, 20

表 3 金銭要求ゲームの利得行列

の確率で選択する行動が混合戦略ナッシュ均衡となる。

3 実験内容

3.1 実験の目的

はじめにでも述べたように、平均値推測ゲームと金銭要求ゲームは、レベル k モデルにもとづく被験者の推論レベルの測定にしばしば使用されるゲームである。したがって、どちらのゲームをもとに推論レベルを測定しても、同じ被験者の推論レベルには違いがないことが望ましい。しかしながら、Georganas et al. (2015) が明らかにしたように、被験者の推論レベルが測定するゲームに依存し、異なるゲーム間での推論レベルの一貫性がみられない可能性がある。

そこで、本研究の目的は、被験者の推論レベルを測定するゲームとして代表的な平均値推測ゲームと金銭要求ゲームを使って、推論レベルを測定するゲームが異なっても、測定される推論レベルに一貫性がみられるかどうかを経済実験によって検証することである。本研究では、青山学院大学の経済学部 2 年生以降を履修対象とするゲーム理論の授業の履修者を被験者として、2017 年 9 月 26 日に経済実験を実施した。今回の実験の被験者の概要を示す表 4 によれば、被験者の多数は 2 年生であることがわかる。今回の実験はゲーム理論の 2 回目の授業で実施したということもあり、被験者の多数を占める 2 年生は、ゲーム理論についての知識はまだほとんど持っていないと思われる。

被験者数	平均値推測ゲームでの グループの構成人数
167 人 (うち 2 年生: 141 人)	16-35 人

表 4 実験の被験者の概要

本論文で検証したい仮説は次のとおりである。

仮説 1. 同一被験者にとって、平均値推測ゲームによる推論レベルと金銭要求ゲームによる推論レベルに違いはない。

特に、非戦略的に行動する推論レベル 0 の被験者、そして、そのプレイヤーに対して最適反応をとる推論レベル 1 の被験者は、平均値推測ゲームと金銭要求ゲームのどちらにおいても行動基準に違いはないため、仮説 1 は次の仮説 2 と仮説 3 を意味する。

仮説 2. 平均値推測ゲームによる推論レベルと金銭要求ゲームによる推論レベルに関して、一方のゲームによる推論レベルがレベル 0 である被験者は、もう一方のゲームによる推論レベルもレベル 0 である。

仮説 3. 平均値推測ゲームによる推論レベルと金銭要求ゲームによる推論レベルに関して、一方のゲームによる推論レベルがレベル 1 である被験者は、もう一方のゲームによる推論レベルもレベル 1 である。

しかしながら、はじめにでも述べたように、平均値推測ゲームと金銭要求ゲームでは推論レベルの分布が異なるために、推論レベル 0 と推論レベル 1 以外については、一貫性を持つ被験者であっても両ゲーム間で推論レベルに違いがでてくる可能性がある。よって、仮説 1 をそのまま検証することはできないため、相対的推論レベルを使って仮説 1 を検証する。

定義 3 (相対的推論レベル). 相対的推論レベルは、最頻値の推論レベルを平均相対的レベル、それよりも低い推論レベルを低い相対的レベル、それよりも高い推論レベルを高い相対的レベルとよぶ。

したがって、仮説 1 は次のように修正される。

仮説 4. 平均値推測ゲームによる推論レベルと金銭要求ゲームによる推論レベルに関して、

- (1) 一方のゲームによる推論レベルが低い相対的レベルである被験者は、もう一方のゲームによる推論レベルも低い相対的レベルである。
- (2) 一方のゲームによる推論レベルが平均相対的レベルである被験者は、もう一方のゲームによる推論レベルも平均相対的レベルである。
- (3) 一方のゲームによる推論レベルが高い相対的レベルである被験者は、もう一方のゲームによる推論レベルも高い相対的レベルである。

本論文では、以上より、仮説 2、仮説 3、仮説 4 を検証していく。

3.2 実験の実施方法

実験では、実験者がスクリーンを使ってゲームのルールを説明した後、被験者は 3 分程度の時間を与えられて意思決定を行った。なお、スクリーンに表示されたゲームのルールは、平均値推測ゲームでは、以下のとおりである。

- (1) 参加者の一人一人が、0 から 100 まで (0 と 100 を含む) の数字をひとつ選んで記入する。
- (2) 実験者が、それらの記入された数字の平均値をグループごとに計算する。
- (3) グループごとに、「推測値」に最も近い数字を記入していた人 (たち) が勝利する (獲得ポイントは、推測値に近い数字を記入しているほど高くなる)。
- (4) 推測値: $\text{平均値} \times \frac{2}{3}$

また、金銭要求ゲームでは、以下のとおりである。

- (1) 参加者は、ランダムに選ばれた相手とペアになり、11 から 20 まで (11 と 20 を含む) の整数を同時にひとつ選択する。
- (2) (1) で自分の選んだ整数は、そのままポイントとして獲得することができる。
- (3) ただし、(1) でペアの相手よりもちょうど 1 だけ小さい整数を選択していた参加者のみ、追加で 20 ポイント獲得することができる。

ゲームの順序は、最初に平均値推測ゲームを行い、次に金銭要求ゲームを行った。なお、平均値推測ゲームおよび金銭要求ゲームでの被験者の選択結果はすべての実験の終了後に公開するようにし、被験者にとっては、選択結果のフィードバックがない状況での意思決定となっている⁴⁾。

被験者による意思決定は、スマートフォンを使ってウェブ上の回答フォームに記入し行われた。なお、被験者間のコミュニケーションは控えるように伝えられている。大教室での教室実験であるため、被験者間の多少のコミュニケーションはあったかもしれないが、被験者数が 167 人と多いため、実験結果に影響を及ぼすことはほとんどないと考えられる。

授業の履修者を被験者としているために、成績の一部を報酬として利用している。具体的には、平均値推測ゲームでは最高 2 点 ($C = 2$) を、金銭要求ゲームでは獲得したポイントの $\frac{1}{10}$ を成績評価の際に期末試験の得点に加算すると実験実施前に伝えている。

4 実験結果

4.1 平均値推測ゲーム

最初に、平均値推測ゲームでの選択結果を、横軸に 0 から 100 の数字、縦軸にその数字の選択割合をとってグラフにしたものを図 1 に、そして、要約統計量を表 5 に示す。また、図 1 の選択結果を、表 1 を使って平均値推測ゲームでの推論レベルに書き換えたものを図 2 に示す。

平均値	中央値	最頻値	最大値	最小値
31.38	33	50	95	0

表 5 要約統計量 (平均値推測ゲーム)

表 5 および図 2 によれば、今回の平均値推測ゲームでは、選択の平均値は 31.38、中央値は 33、そして、推論レベルの最頻値はレベル 1 である。一方、Nagel (1995) による平均値推測ゲームでは、選択の平均値は 36.73、中央値は 33、

4) なお、平均値推測ゲームの前に別のゲームを実施しているが、そのゲームの選択結果の公開もすべての実験の終了後に行っている。

推論レベルの一貫性

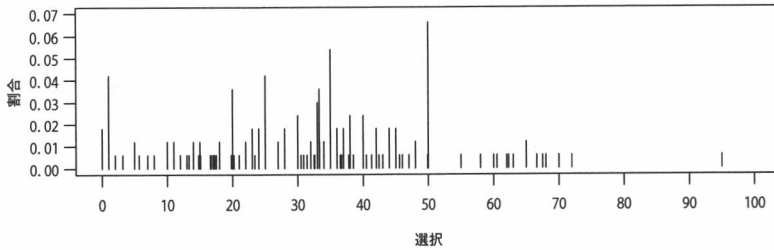


図1 選択別の被験者の割合 (平均値推測ゲーム)

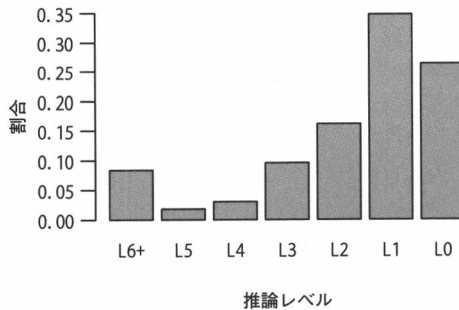


図2 推論レベル別の被験者の割合 (平均値推測ゲーム)

推論レベルの最頻値はレベル1であった。したがって、今回の平均値推測ゲームの結果は、Nagel (1995) などの先行研究での実験結果と違いはないと考えられる。

4.2 金銭要求ゲーム

次に、金銭要求ゲームでの選択結果を、横軸に11から20の数字、縦軸にその数字の選択割合をとってグラフにしたものを図3に、そして、要約統計量を表6に示す。また、図3の選択結果を、表2を使って金銭要求ゲームでの推論レベルに書き換えたものを図4に示す。

平均値	中央値	最頻値	最大値	最小値
17.88	18	18	20	11

表6 要約統計量 (金銭要求ゲーム)

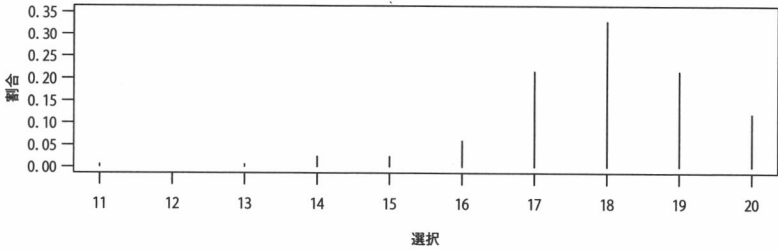


図 3 選択別の被験者の割合 (金銭要求ゲーム)

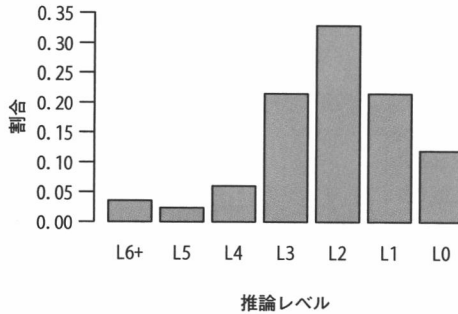


図 4 推論レベル別の被験者の割合 (金銭要求ゲーム)

表 6 および図 4 によれば、今回の金銭要求ゲームでは、選択の平均値は 17.88, 中央値は 18, そして、推論レベルの最頻値はレベル 2 である。一方, Arad and Rubinstein (2012) による金銭要求ゲームでは、選択の平均値は 17.15, 中央値は 17, 推論レベルの最頻値はレベル 3 であった。したがって、今回の金銭要求ゲームの結果は、推論レベルの最頻値が違うこともあり, Arad and Rubinstein (2012) の結果とは少し異なるものであるようにもみえる。そこで、今回の結果が例外的なものであるのかを確認するために、金銭要求ゲームに関するいくつかの実験結果を表 7 にまとめた。すると、Arad and Rubinstein (2012) 以外による金銭要求ゲームでは、選択の中央値は 18, 推論レベルの最頻値はレベル 2 であることがわかる。したがって、今回の金銭要求ゲームの結果は、Arad and Rubinstein (2012) の結果とは少し異なるが、先行研究での実験結果と違いはないと考えられる。

推論レベルの一貫性

	中央値	推論レベルの最頻値
Arad and Rubinstein (2012)	17	レベル 3
Lindner and Sutter (2013)	18	レベル 2
Goeree et al. (2017)	18	レベル 2

表 7 金銭要求ゲームの主な実験結果

5 推論レベルの一貫性

この章では、4 章で紹介した実験結果を使って、平均値推測ゲームと金銭要求ゲームによる被験者の推論レベルの測定結果に一貫性があるのかどうかを検証していく。まずは、同一被験者の両ゲーム間での推論レベルの変化を示すものとして、横軸に平均値推測ゲームでの推論レベル、縦軸に金銭要求ゲームでの推論レベルをとってグラフにしたものを図 5 に示す。また、同一被験者の両ゲーム間での推論レベルの変化量を測るものとして、推論レベル差を

$$|\text{平均値推測ゲームによる推論レベル} - \text{金銭要求ゲームによる推論レベル}|$$

で定義し、推論レベル差別に被験者の割合をまとめたものを表 8 に示す。

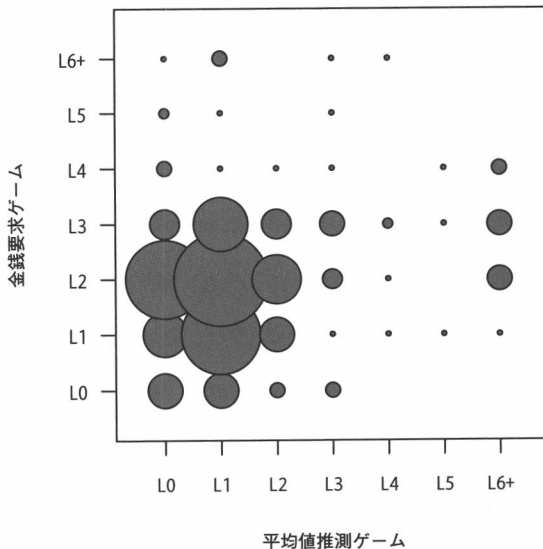


図 5 両ゲーム間での推論レベルの変化

推論レベル差	人数	割合
0	38 人	22.75%
1	56 人	33.53%
2	39 人	23.35%
3	17 人	10.18%
4	10 人	5.99%
5	6 人	3.59%
6	1 人	0.60%

表 8 推論レベル差別の被験者の割合

図 5 および表 8 によれば、平均値推測ゲームと金銭要求ゲームの間で推論レベルが一貫している被験者（推論レベル差が 0 である被験者）の割合は、22.75% のみであることがわかる。この値は、先行研究である Dittrich and Leipold (2014) の割合 80.28% と比べてかなり低くなっている。

このように、今回の実験結果からは、同一被験者であったとしても、平均値推測ゲームでの推論レベルと金銭要求ゲームでの推論レベルには、一貫性がなく大きな違いがあるように見える。このことは、推論レベル測定ゲーム別に両ゲーム間での推論レベルの変化を示す図 6 からも確認できる⁵⁾。実際、推論レベル測定ゲーム別に推論レベルの一貫している被験者の割合を示す表 9 によれば、最高で、平均値推測ゲームでの推論レベルがレベル 2 である被験者の 37.04%、および、金銭要求ゲームでの推論レベルがレベル 1 である被験者の 44.44% が推論レベルの一貫性を保っているが、いずれの値も 50% 未満となっている。

以上より、仮説 1 に関しては次のことがわかった。

結果 1 (仮説 1). 平均値推測ゲームと金銭要求ゲーム間で、被験者の推論レベルに一貫性はない。

次節からは、これらの実験結果をもとに、仮説 2、仮説 3、仮説 4 の検証を行っていく。

5) 平均値推測ゲームでの推論レベルを基準に、金銭要求ゲームでの推論レベルの変化を示すのが図 6 左、金銭要求ゲームでの推論レベルを基準に、平均値推測ゲームでの推論レベルの変化を示すのが図 6 右である。

推論レベルの一貫性

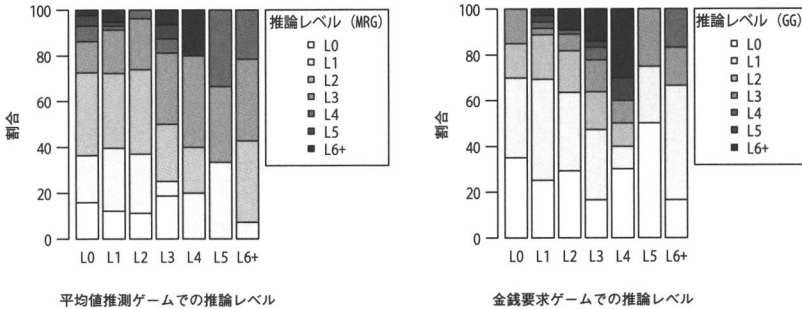


図6 両ゲーム間での推論レベルの変化(推論レベル測定ゲーム別)

推論レベル	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
レベル0	15.91% (7/44)	35.00% (7/20)
レベル1	27.59% (16/58)	44.44% (16/36)
レベル2	37.04% (10/27)	18.18% (10/55)
レベル3	31.25% (5/16)	13.89% (5/36)
レベル4	0.00% (0/5)	0.00% (0/10)
レベル5	0.00% (0/3)	0.00% (0/4)
レベル6+	0.00% (0/14)	0.00% (0/6)

表9 推論レベルの一貫している被験者の割合(推論レベル測定ゲーム別)

5.1 推論レベル0の一貫性

まずは、仮説2を検証するために、推論レベル0である被験者についてみていこう。推論レベル0である被験者は、平均値推測ゲームによる推論レベルがレベル0である被験者と金銭要求ゲームによる推論レベルがレベル0である被験者の二通り存在する。そのため、推論レベル0である被験者について、測定ゲーム別に推論レベル差を計算し、それをまとめたものを表10に示す。

表10または表9によれば、両ゲーム間で推論レベルの一貫性を保つ被験者(推論レベル差が0である被験者)の割合は、平均値推測ゲームによる推論レベルでは15.91%、金銭要求ゲームによる推論レベルでは35.00%である。したがって、推論レベル0の被験者については、両ゲーム間での推論レベルの一貫性はみられないことがわかった。

結果2(仮説2). 推論レベル0の被験者については、平均値推測ゲームと金銭要

推論レベル差	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
0	15.91% (7/44)	35.00% (7/20)
1	20.45% (9/44)	35.00% (7/20)
2	36.36% (16/44)	15.00% (3/20)
3	13.64% (6/44)	15.00% (3/20)
4	6.82% (3/44)	0.00% (0/20)
5	4.55% (2/44)	0.00% (0/20)
6	2.27% (1/44)	0.00% (0/20)

表 10 推論レベル差別の被験者の割合 (推論レベル 0 の被験者, 推論レベル測定ゲーム別)
求ゲーム間で, 被験者の推論レベルに一貫性はない。

両ゲーム間での一貫性を保つ被験者の割合は, 金銭要求ゲームによる推論レベル (35.00%) の方が平均値推測ゲームによる推論レベル (15.91%) よりも多い。そこで, 推論レベルの一貫性を持つ被験者の割合は, 金銭要求ゲームによる推論レベルの方が, 平均値推測ゲームによる推論レベルよりも多いかどうかを検証してみよう。すると, フィッシャーの正確確率検定を使って検証すると,

- 推論レベルが一貫している被験者の割合は, 金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも多い ($p = 0.08504$) ことがわかった。

したがって, 仮説 2 に関係することとして, 次のことがいえるだろう。

結果 3 (仮説 2). 推論レベル 0 の被験者については, 金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも推論レベルが一貫している可能性が高い。

5.2 推論レベル 1 の一貫性

次に, 仮説 3 の検証を行っていく。推論レベル 1 である被験者も, 推論レベル 0 である被験者と同様に, 推論レベルを測定するゲーム別に二通り存在する。そこで, 推論レベル 1 である被験者について, 測定ゲーム別に推論レベル差を計算し, それをまとめたものを表 11 に示す。

表 11 または表 9 によれば, 両ゲーム間で推論レベルの一貫性を保つ被験者

推論レベルの一貫性

推論レベル差	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
0	27.59% (16/58)	44.44% (16/36)
1	44.83% (26/58)	44.44% (16/36)
2	18.97% (11/58)	2.78% (1/36)
3	1.72% (1/58)	2.78% (1/36)
4	1.72% (1/58)	2.78% (1/36)
5	5.17% (3/58)	2.78% (1/36)

表 11 推論レベル差別の被験者の割合 (推論レベル 1 の被験者, 推論レベル測定ゲーム別) の割合は, 平均値推測ゲームによる推論レベルでは 27.59%, 金銭要求ゲームによる推論レベルでは 44.44% である。したがって, 推論レベル 1 の被験者についても, 両ゲーム間での推論レベルの一貫性はみられないことがわかった。

結果 4 (仮説 3). 推論レベル 1 の被験者については, 平均値推測ゲームと金銭要求ゲーム間で, 被験者の推論レベルに一貫性はない。

両ゲーム間での一貫性を保つ被験者の割合は, 金銭要求ゲームによる推論レベル (44.44%) の方が平均値推測ゲームによる推論レベル (27.59%) よりも多い。そこで, 推論レベルの一貫性を持つ被験者の割合は, 金銭要求ゲームによる推論レベルの方が, 平均値推測ゲームによる推論レベルよりも多いかどうかを検証してみよう。そこで, フィッシャーの正確確率検定を使って検証すると,

- 推論レベルの一貫している被験者の割合は, 金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも多い ($p = 0.07367$)

ことがわかった。

したがって, 仮説 3 に関係することとして, 次のことがいえるだろう。

結果 5 (仮説 3). 推論レベル 1 の被験者については, 金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも推論レベルが一貫している可能性が高い。

5.3 相対的推論レベルの一貫性

最後に、仮説 4 を検証するために、被験者の相対的推論レベルについてみていこう。同一の被験者の平均値推測ゲームと金銭要求ゲーム間での相対的推論レベルの変化を示すものとして、横軸に平均値推測ゲームでの相対的推論レベル、縦軸に金銭要求ゲームでの相対的推論レベルをとってグラフにしたものを図 7 に示す。また、同一被験者の両ゲーム間での相対的推論レベルの変化量を測るものとして、相対的推論レベル差を

$|\text{平均値推測ゲームによる相対的推論レベル} - \text{金銭要求ゲームによる相対的推論レベル}|$ で定義し、相対的推論レベル差ごとの被験者の割合をまとめたものを表 12 に示す⁶⁾。

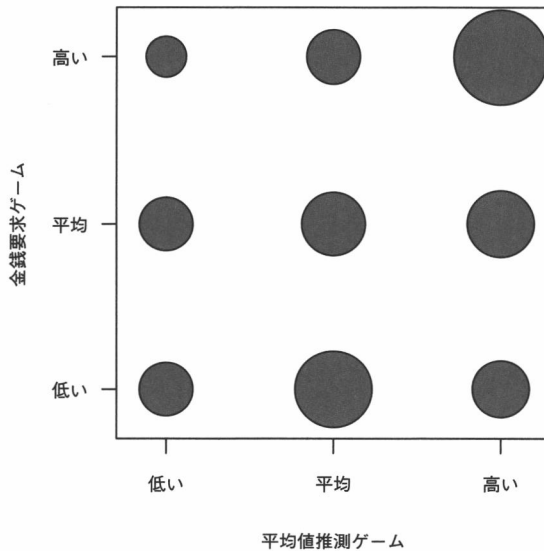


図 7 両ゲーム間での相対的推論レベルの変化

図 7 および表 12 によれば、平均値推測ゲームと金銭要求ゲームの間で相対的推論レベルが一貫している被験者の割合は、37.72% であることがわかる。こ

6) 相対的推論レベル差は、低い相対的推論レベルの値を 1、平均相対的推論レベルの値を 2、高い相対的推論レベルの値を 3 として計算している。

推論レベルの一貫性

相対的推論レベル差	人数	割合
0	63 人	37.72%
1	75 人	44.91%
2	29 人	17.37%

表 12 相対的推論レベル差別の被験者の割合

の値は、絶対的推論レベルが一貫している被験者の割合 22.75% よりも高くなっているが、先行研究である Dittrich and Leipold (2014) の割合 80.28% と比べてかなり低くなっている。

このように、今回の実験結果からは、平均値推測ゲームでの相対的推論レベルと金銭要求ゲームでの相対的推論レベルには、絶対的推論レベルと同様に、一貫性がなく大きな違いがあるようにみえる。このことは、推論レベル測定ゲーム別に両ゲーム間での相対的推論レベルの変化を示す図 8 から確認できる⁷⁾。実際、推論レベル測定ゲーム別に相対的推論レベルの一貫している被験者の割合を示す表 13 によれば、最高で、平均値推測ゲームでの推論レベルが高い相対的レベルである被験者の 43.08%、および、金銭要求ゲームでの推論レベルが高い相対的レベルである被験者の 50.00% が相対的推論レベルの一貫性を保っているが、いずれの値も 50% 以下となっている。

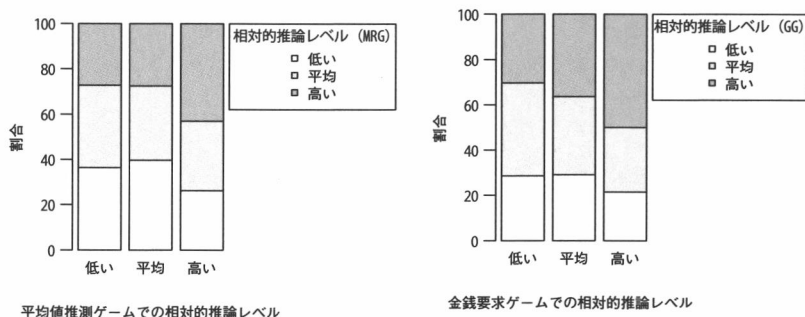


図 8 両ゲーム間での相対的推論レベルの変化 (推論レベル測定ゲーム別)

7) 平均値推測ゲームでの相対的推論レベルを基準に、金銭要求ゲームでの相対的推論レベルの変化を示すのが図 8 左、金銭要求ゲームでの相対的推論レベルを基準に、平均値推測ゲームでの相対的推論レベルの変化を示すのが図 8 右である。

相対的推論レベル	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
低い相対的レベル	36.36% (16/44)	28.57% (16/56)
平均相対的レベル	32.76% (19/58)	34.55% (19/55)
高い相対的レベル	43.08% (28/65)	50.00% (28/56)

表 13 相対的推論レベルの一貫している被験者の割合 (推論レベル測定ゲーム別)

相対的推論レベルについても、絶対的推論レベルと同様に、測定するゲーム別に二通り存在する。そこで、低い・平均・高い相対的推論レベルである被験者について、測定ゲーム別に推論レベル差を計算し、それらをまとめたものを表 14・15・16 に示す。

推論レベル差	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
0	36.36% (16/44)	28.57% (16/56)
1	36.36% (16/44)	41.07% (23/56)
2	27.27% (12/44)	30.36% (17/56)

表 14 推論レベル差別の被験者の割合 (低い相対的推論レベルの被験者，推論レベル測定ゲーム別)

推論レベル差	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
0	32.76% (19/58)	34.55% (19/55)
1	67.24% (39/58)	65.45% (36/55)

表 15 推論レベル差別の被験者の割合 (平均相対的推論レベルの被験者，推論レベル測定ゲーム別)

表 14・15・16 または表 13 によれば、両ゲーム間で相対的推論レベルの一貫性を保つ被験者の割合は、低い相対的推論レベルの被験者については、平均値推測ゲームによる推論レベルでは 36.36%，金銭要求ゲームによる推論レベルでは 28.57%，平均相対的推論レベルの被験者については、平均値推測ゲームによる推論レベルでは 32.76%，金銭要求ゲームによる推論レベルでは 34.55%，高い相対的推論レベルの被験者については、平均値推測ゲームによる推論レベルでは 43.08%，金銭要求ゲームによる推論レベルでは 50.00% である。したがって、どの相対的推論レベルの被験者についても、両ゲーム間での相対的推論レ

推論レベルの一貫性

推論レベル差	推論レベル測定ゲーム	
	平均値推測ゲーム	金銭要求ゲーム
0	43.08% (28/65)	50.00% (28/56)
1	30.77% (20/65)	28.57% (16/56)
2	26.15% (17/65)	21.43% (12/56)

表 16 推論レベル差別の被験者の割合 (高い相対的推論レベルの被験者、推論レベル測定ゲーム別)

ベルの一貫性はみられないことがわかった。

結果 6 (仮説 4). 低い相対的推論レベルの被験者については、平均値推測ゲームと金銭要求ゲーム間で、被験者の相対的推論レベルに一貫性はない。

結果 7 (仮説 4). 平均相対的推論レベルの被験者については、平均値推測ゲームと金銭要求ゲーム間で、被験者の相対的推論レベルに一貫性はない。

結果 8 (仮説 4). 高い相対的推論レベルの被験者については、平均値推測ゲームと金銭要求ゲーム間で、被験者の相対的推論レベルに一貫性はない。

両ゲーム間での一貫性を保つ被験者の割合は、低い相対的推論レベルの被験者については、平均値推測ゲームによる推論レベル (36.36%) の方が金銭要求ゲームによる推論レベル (28.57%) よりも多く、平均相対的推論レベルの被験者については、金銭要求ゲームによる推論レベル (34.55%) の方が平均値推測ゲームによる推論レベル (32.76%) よりも多く、高い相対的推論レベルの被験者については、金銭要求ゲームによる推論レベル (50.00%) の方が平均値推測ゲームによる推論レベル (43.08%) よりも多い。そこで、相対的推論レベルの一貫性を持つ被験者の割合は、どちらのゲームによる推論レベルの方がより多いのかについて検証してみよう。すると、フィッシャーの正確確率検定を使って検証すると、

- 低い相対的推論レベルの被験者については、相対的推論レベルの一貫している被験者の割合は、平均値推測ゲームによる推論レベルの方が金銭要求ゲームによる推論レベルより多いとはいえない
- 平均相対的推論レベルの被験者については、相対的推論レベルの一貫し

ている被験者の割合は、金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも多いとはいえない

- 高い相対的推論レベルの被験者については、相対的推論レベルの一貫している被験者の割合は、金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも多いとはいえない

ことがわかった。

表14・15・16によれば、相対的推論レベル差が0である被験者の割合は、高い相対的推論レベルの被験者についてのみ、平均値推測ゲームによる推論レベル・金銭要求ゲームによる推論レベルのいずれにおいても最大になっている。そこで、相対的推論レベル差が0である被験者の割合が、どちらのゲームにおいても最大になっているのかどうかについて検証してみよう。すると、フィッシャーの正確確率検定を行うと、

- 平均値推測ゲームにおいて高い相対的推論レベルの被験者については、相対的推論レベル差が0である被験者の割合は、相対的推論レベル差が1である被験者の割合よりも多いとはいえないが ($p = 0.1016$)、相対的推論レベル差が2である被験者の割合よりも多いといえる ($p = 0.03233$)
- 金銭要求ゲームにおいて高い相対的推論レベルの被験者については、相対的推論レベル差が0である被験者の割合は、相対的推論レベル差が1である被験者の割合よりも多く ($p = 0.01638$)、相対的推論レベル差が2である被験者の割合よりも多い ($p = 0.001431$)

ことがわかった。

以上より、仮説4に関係することとして、次のことがいえるだろう。

結果9 (仮説4). 金銭要求ゲームで高い相対的推論レベルの被験者は、相対的推論レベルが一貫している可能性が、相対的推論レベルが1だけ変化する可能性、相対的推論レベルが2だけ変化する可能性よりも高い。

6 おわりに

本論文では、被験者の推論レベルが測定するゲームの種類に依存せずに、被験者にとって一貫性を持つものであるのかどうかを明らかにすることを試みた。そのために、平均値推測ゲームと金銭要求ゲームをもとに被験者の推論レベルの測定を行い、同一被験者の推論レベルが両ゲーム間で一貫しているかどうかを検証した。その結果、平均値推測ゲームと金銭要求ゲーム間での推論レベルの一貫性はみられないことがわかった。

次に、絶対的推論レベルではなく相対的推論レベルに着目し、同一被験者の相対的推論レベルが両ゲーム間で一貫しているかどうかを検証した。その結果、相対的推論レベルに関しても、絶対的推論レベルと同様に、平均値推測ゲームと金銭要求ゲーム間での一貫性はみられないことがわかった。

以上より、被験者の選択理由ではなく選択結果を使って被験者の推論レベルの測定を行い Dittrich and Leipold (2014) の結果を再検証したところ、平均値推測ゲームと金銭要求ゲーム間で、絶対的推論レベル・相対的推論レベルの両方において、推論レベルの一貫性はみられないことがわかった。しかしながら、推論レベル 0 および推論レベル 1 である被験者に関しては、推論レベルが一貫している可能性は、金銭要求ゲームによる推論レベルの方が平均値推測ゲームによる推論レベルよりも高いことがわかった。さらに、金銭要求ゲームにおいて高い相対的推論レベルを持つ被験者に関しては、相対的推論レベルが一貫している可能性が、相対的推論レベルが 1 だけ変化する可能性、相対的推論レベルが 2 だけ変化する可能性よりも高いこともわかった。これらの意味において、金銭要求ゲームによる推論レベルの方が、平均値推測ゲームによる推論レベルよりも一貫性を持つといえるだろう。

本研究では、ゲーム理論の授業の履修者を被験者として実験を実施している。そのため、本研究でみられた結果が、別の被験者でも成立するのかどうかを検証することは今後の課題である。

参考文献

- Alaoui, L. and A. Penta (2016): "Endogenous Depth of Reasoning," *Review of Economic Studies*, Vol. 83 (4), 1297-1333, doi:10.1093/restud/rdv052.
- Allred, S., S. Duffy, and J. Smith (2016): "Cognitive Load and Strategic Sophistication," *Journal of Economic Behavior and Organization*, Vol. 125, 162-178, doi:10.1016/j.jebo.2016.02.006.
- Arad, A. and A. Rubinstein (2012): "The 11-20 Money Request Game: A Level- k Reasoning Study," *American Economic Review*, Vol. 102 (7), 3561-3573, <http://www.jstor.org/stable/41724645>.
- Bosch-Domènech, A. and R. Nagel (1997): "El Juego de Adivinar el Número X: Una Explicación y la Proclamación del Vencedor," *Expansión*, 42-43, June 16.
- Bühren, C. and B. Frank (2012): "Chess Players' Performance beyond 64 Squares: A Case Study on the Limitations of Cognitive Abilities Transfer," *Talent Development and Excellence*, Vol. 4 (2), 157-169, http://www.iratde.org/images/stories/journal/2012-2/2012-2-07buhren_frank.pdf.
- Camerer, C. F. (1997): "Progress in Behavioral Game Theory," *Journal of Economic Perspectives*, Vol. 11 (4), 167-188, doi:10.1257/jep.11.4.167.
- Costa-Gomes, M. A. and V. P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, Vol. 96 (5), 1737-1768, doi:10.1111/1468-0262.00239.
- Dittrich, M. and K. Leipold (2014): "Gender Differences in Strategic Reasoning," CESifo Working Paper 4763, Ludwig-Maximilians University, 1-18, http://www.cesifo-group.de/DocDL/cesifo1_wp4763.pdf.
- Duffy, J. and R. Nagel (1997): "On the Robustness of Behaviour in Experimental 'Beauty Contest' Games," *Economic Journal*, Vol. 107 (445), 1684-1700, doi:10.1111/j.1468-0297.1997.tb00075.x.
- Georganas, S., P. J. Healy, and R. A. Weber (2015): "On the Persistence of Strategic Sophistication," *Journal of Economic Theory*, Vol. 159 (Part A), 369-400, doi:10.1016/j.jet.2015.07.012.
- Goeree, J. K., P. Louis, and J. Zhang (2017): "Noisy Introspection in the 11-20 Game," *Economic Journal*, 1-22, doi:10.1111/eoj.12479.
- Lindner, F. and M. Sutter (2013): "Level- k Reasoning and Time Pressure in the 11-20 Money Request Game," *Economics Letters*, Vol. 120 (3), 542-545, doi:10.1016/j.econlet.2013.06.005.
- Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, Vol. 85 (5), 1313-1326, <http://www.jstor.org/stable/2950991>.
- (2008): "Experimental Beauty Contest Games: Levels of Reasoning and Convergence to Equilibrium," in Plott, C. R. and V. L. Smith eds. *Handbook of Experimental Economics Results*, Vol. 1: Elsevier B.V., 391-410, doi:10.1016/S1574-0722(07)00045-5.
- Selten, R. and R. Nagel (1998): "Das Zahlenwahlspiel—Ergebnisse und Hintergrund," *Spektrum der Wissenschaft*, 16-22, February.
- de Sousa, J., G. Hollard, and A. Terracol (2014): "Non-Strategic Players are the Rule rather than the

推論レベルの一貫性

Exception,” Mimeo, 1–30, <https://sites.google.com/site/webpagehollard/Chess-levelK.pdf>.

Thaler, R. (1997): “Giving Markets a Human Dimension,” *Financial Times*, 2–5, June 16.

水上英貴 (2015): 「平均値推測ゲーム実験—青山学院での教室実験結果—」, 『青山経済論集』, 3 (67), 111–138.