

ゲーミフィケーションのルールがレビュー及びコーディングに及ぼす影響の分析

吉上 康平^{†1} 林 大志^{†1} 角田 雅照^{†1,a)} 上野 秀剛^{†2,b)}

開発の効率を高めるためのアプローチとして、ゲーミフィケーションを適用することが試みられており、それにより、開発者のモチベーションを高め、効率が改善することが期待されている。ゲーミフィケーションとは、ある作業に娛樂性を付与することであり、それにより作業意欲を向上させることを目的としている。ゲーミフィケーションをソフトウェア開発に適用する場合、様々なルール設定が考えられるが、そのルールが作業効率にどのように影響するかは明らかでない。また、個人の嗜好などにより、ゲーミフィケーションのルールの効果が異なる可能性がある。そこで本研究では、ゲーミフィケーションのルールの差異が作業の成果に影響するかどうか、また、ゲーミフィケーションのルールに対する評価が個人により異なるかどうかを分析した。

キーワード： 人的要因、モチベーション、時間割引率

1. はじめに

昨今、ソフトウェア開発は大規模化しており、ソフトウェア開発の効率をいかにして高めるかが、非常に重要な課題のひとつとなっている。開発の効率を高めるためのアプローチとして、ゲーミフィケーションを適用することが試みられており、それにより、開発者のモチベーションを高め、効率が改善することが期待されている。ゲーミフィケーションとは、ある作業に娛樂性を付与することであり、それにより作業意欲を向上させることを目的としている。ゲーミフィケーションでよく用いられる要素として、記録の競争や、架空のトロフィーやバッジの収集などがあげられ、前者では競争、後者では収集に焦点を置き、モチベーションを高めるアプローチとしている。

ゲーミフィケーションをソフトウェア開発に適用する場合、様々なルール設定が考えられるが、そのルールが作業効率にどのように影響するかは明らかでない。例えば、ゲーミフィケーションのルールとして、作業時間に制限を設けた場合、作業時間が早まる可能性があるが、逆に心理的に焦りが生じることにより、作業の成果に何らかの悪影響を及ぼす可能性がある。具体的には、コードレビューにおいて時間制限のルールを適用し場合、レビュー時間は短くなる可能性があるが、逆にレビュー指摘数(発見欠陥数)が低下する可能性がある。

また、個人の嗜好などにより、ゲーミフィケーションのルールの効果が異なる可能性がある。例えば、焦りが生じやすい性格を持つ開発者に対し、プログラミング時に作業時間とともにスコアが減少するルールを適用した場合、逆に心理的プレッシャーから作業効率が低下することも考え

うる。

そこで本研究では、ゲーミフィケーションのルールの差異が作業の成果に影響するかどうか、また、ゲーミフィケーションのルールに対する評価が個人により異なるかどうかを分析する。そのために、実験において2つのタスクを被験者に与え、その結果を分析した。ひとつのタスクは、コードレビューであり、もうひとつのタスクはプログラミングである。個人の嗜好に関しては、アンケートによる評価にもとづいて評価するとともに、行動経済学で定義されている時間割引率との関係についても分析した。

2. ゲーミフィケーション

ゲーミフィケーションとは、作業に娛樂性を付加することで、ソフトウェア開発に限らず、近年非常に注目されているアプローチである[4][9]。ゲーミフィケーションの目的は、作業効率や持続性の向上である。例えば、ある作業を完了した際に何らかのアイテムやバッジなどが入手できるようにすることがある。これは、アイテムを収集したいという欲求を刺激することにより、作業に対するモチベーションを高めている。また、作業に対して得点(スコア)を算出することも多い。例えば、ある作業を完了するごとにスコアが加点され、過去の自分、もしくは他人との競争を可能にすることで作業効率を上げることを狙いとしている。

ソフトウェア開発において、ゲーミフィケーションを適用する試みはいくつか行われている[2][5][6][8][10]。ただし、我々の知る限り、ゲーミフィケーションのルールが作業結果にどのように影響するか、また、個人の嗜好とゲーミフィケーションとの関係はどのようなものかについては明らかにされていない。

3. 時間割引率

時間割引率は行動経済学で定義されている概念である。行動経済学とは、人が必ずしも合理的に活動しないことを

†1 近畿大学
Kindai University, Japan

†2 奈良工業高等専門学校
National Institute of Technology, Nara College, Japan
a) tsunoda@info.kindai.ac.jp
b) uwano@info.nara-k.ac.jp

考慮して経済活動を分析する学問である。従来の経済学では、1年後の資産価値は利子によってのみ決まる。これに対し行動経済学の時間割引率では、各人の心理によって(資産)価値が異なると考える。例えば、現在3,000円受け取ることと、1年後に4,000円受け取ることのどちらかを選択する場合、個人により判断が異なることが実験により示されている。時間割引率の高い人の場合、1年後の金額が大きい場合のみ、低い人の場合、1年後の金額が小さくても後者を選ぶ。

本研究では、時間割引率がゲーミフィケーションのルールに何らかの関係があるかどうかを分析する。時間割引率と、その他の性格(ビックファイブなど)は関連があることが指摘されている[3]。性格に関する評価指標の中で、ゲーミフィケーションの得点と時間割引率が最も関連が強いと仮定し、分析に用いた。時間割引率の求め方については、文献[3]などで示されている方法に従った。

4. 実験

実験では、被験者にレビュー、コーディングのタスクを順番に行ってもらい、コーディングのタスク直後にアンケートを実施した。前者ではコードの誤りを指摘してもらい、後者では仕様に従ってコーディングをしてもらった。レビューとコーディングのタスクの間は1週間以上空いている。被験者には、実験後に自分の順位を知らせる(ただし他の被験者には結果を開示しない)ことを事前に伝え、競争心が高まる工夫を行った。被験者は情報科学を専攻する学部生であり、レビューのタスクの被験者数は13人、コーディングのタスクは14人である。両方のタスクにおいて、被験者はコーディングのタスク1名を除いて同じである。

4.1 レビュータスク

レビューのタスクでは、プログラムの仕様書(誤りなし)とソースコード(誤りあり)を与え、誤りの箇所と修正方法を記述してもらった。ソースコードは被験者全員が理解しているJavaで記述されている。

- レビュー対象は文献[7]で示されている以下の2つとした。
- レビュー対象A: 数字の合計を計算して出力するプログラム。設計書が100字、プログラムが200行程度のもの。誤りが5つ含まれる。
 - レビュー対象B: ファイルを読み込んで目的のデータを検索するプログラム。設計書が600字、プログラムが130行程度。誤りが8つ含まれる。

被験者は最初に対象Aをレビューし、次に対象Bを続けてレビューした。

ゲーミフィケーションのルールは以下の2種類を用意した。

- ルール α : 誤りを指摘した数のみで得点が加算される。予備分析に基づき、レビュー対象Aの得点は200点、レビュー対象Bの得点は250点とした。

- ルール β : 指摘数に基づく点数に加えて、最初に持ち点があり、レビュー時間が5秒経過するごとに5点が減点される。レビュー対象Aの持ち点は1200点、レビュー対象Bの得点は3600点とし、前述2つの合計が表示される。すなわち、時間が経過すると点数が減少するが、指摘するとその点数が回復することになる。レビュー時には指摘内容が正しくても正しくない場合でも得点が加算されることとし、集計時に誤った指摘については加点しないこととした。この集計方法は被験者に事前に伝えた。

ルール β を適用したレビューでは、目標値を提示してモチベーションを高めるために、平均的な得点(予備実験に基づく)。対象Aでは1,000点、対象Bでは3,120点とした)を被験者に示した。ルール α の場合では、目標値は潜在的なバグ数の手掛かりとなるため提示しなかった。

レビュー対象の違いが実験結果に影響することを避けるため、被験者を2つのグループに分けてレビュー対象とルールの組み合わせを変えた。一方のグループ(被験者7人)はルール α を適用してレビュー対象Aを、ルール β を適用してレビュー対象Bをレビューしてもらった。他方のグループ(被験者6人)では逆にルール β を適用してレビュー対象Aを、ルール α を適用してレビュー対象Bをレビューしてもらった。

4.2 コーディングタスク

コーディングのタスクでは、与えられた仕様に基づき、プログラミングをしてもらった。用いるプログラミング言語は、被験者全員が理解しているJavaとし、Eclipse上でコーディングしてもらい、仕様に書いてある実行結果と一致すればコーディング完了とした。

与えた仕様は以下の2つであり、比較的容易なものをWebサイト(AIZU ONLINE JUDGE[1])から選んだ。

- 仕様A: 九九を出力して終了する。
- 仕様B: 10個の値を読み込み、大きい順に3つの値を出力する。

被験者は最初に仕様Aに基づいてコーディングし、次に仕様Bにもとづいてコーディングをした。

ゲーミフィケーションのルールは以下の2種類を用意した。

- ルール γ : スコア(回答時間)を計算するがリアルタイムでは示さず、タスク終了後に順位とともに示す。
- ルール δ : スコア(回答時間)をリアルタイムで示し、かつレビュータスクのスコアとコーディングタスクのスコアを合算したものをタスク後に示す。

仕様の違いが影響することを避けるために、被験者を2つのグループに分け、仕様とルールの組み合わせを変えて実験を行った。一方のグループ(被験者5人)では仕様Aとルール γ 、仕様Bとルール δ を組み合わせて実験を行った。他方のグループ(被験者9人)では、逆に仕様Aにル

ール δ , 仕様 B にルール γ を適用し, 実験を行った.

4.3 アンケート

コーディングタスク実施直後に, 以下のアンケートを実施した. 各項目について, 主観に基づき 5 段階 (1 が最も当てはまらない, 5 が最も当てはまる) で評価してもらった.

- Q1: ゲーミフィケーション(ランキング)により, 作業効率ややる気に差が出たと感じたか?
- Q2: スコアが出る場合と出ない場合で, 作業効率ややる気に差が出たと感じたか?
- Q3: 前回の得点に加算する場合としない場合で, 作業効率ややる気に差が出たと感じたか?

5. 結果

5.1 リサーチクエスチョンの設定

分析の目的を明確にするために, リサーチクエスチョンを設定した. RQ1 はレビューに関する項目, RQ2 はコーディングに関する項目, RQ3 はゲーミフィケーションに対する主観的評価に関する項目, RQ4 はゲーミフィケーションの好みに影響する要因に関する項目である.

- RQ1-1: ゲーミフィケーションにより, レビュー指摘数は増加するのか?
- RQ1-2: ゲーミフィケーションのルールにより, レビュー時間は短縮されるのか?
- RQ1-3: ゲーミフィケーションのルールにより, レビュー結果が悪影響を受けるのか?
- RQ2-1: ゲーミフィケーションにより, コーディング時間は短縮されるのか?
- RQ2-2: ゲーミフィケーションのルールにより, コーディング時間に与える影響に違いはあるのか?
- RQ3-1: ゲーミフィケーションは主観的にどのように評価されているか?
- RQ3-2: ゲーミフィケーションのルールに対して, 個人の評価差は大きいのか?
- RQ4: ゲーミフィケーションの好みに影響する要因は何か?

5.2 コードレビュー

分析では, ゲーミフィケーションを好むほどその効果が大きいと仮定し, 質問項目 Q1 とレビュー指摘数などの相関係数を算出した.

ゲーミフィケーションの効果: まず, Q1 と指摘数との関係について分析した. Q1 とレビュー指摘数との相関係数を表 1 に示す. Q1 とレビュー対象 B, レビュー対象 A, B の結果の平均との相関係数がそれぞれ 0.3 を超えていた. レビュー対象 A の場合, Q1 と指摘数の相関がほとんどなかったが, これは, レビュー対象 A の大きさが小さく, 指摘数にあまり差が出なかつたためであると考えられる. のことから, ゲーミフィケーションはレビュー指摘数の増

表 1 Q1 とレビュー指摘数との相関係数

レビュー対象 A	レビュー対象 B	平均
-0.05	0.34	0.33

表 2 Q1 とレビュー結果との相関係数

(a) レビュー時間		
レビュー対象 A	レビュー対象 B	
ルール α	-0.66	0.65
ルール β	-0.28	-0.33
(b) レビュー指摘数		
レビュー対象 A	レビュー対象 B	
ルール α	-0.35	0.27
ルール β	-0.10	0.48

表 3 Q1 と正規化したレビュー結果との相関係数

ルール β 適用対象	レビュー時間	レビュー指摘数
レビュー対象 A	0.72	0.24
レビュー対象 B	0.63	0.51

表 4 ルール β とレビュー効率との相関係数

ルール β 適用対象	レビュー 対象 A	レビュー 対象 B
レビュー対象 A	0.13	0.18
レビュー対象 B	0.18	0.30

加に一定の効果があると考えられる.

ゲーミフィケーションのルールの影響: 次に, ゲーミフィケーションのルールの違いが, レビュー結果に影響するかを分析するため, ルールとレビュー対象の組み合わせによってデータを層別した. 相関係数を表 2 に示す. レビュー対象 B の場合, ルール α の相関係数は小さいものの, どちらのルールの場合でも, レビュー指摘数と Q1 とに正の相関があった. また, ルール α の場合, レビュー時間と Q1 は正の相関, ルール β の場合では負の相関が見られた. レビュー対象 A の場合, ルール α , β どちらの場合でも, Q とレビュー指摘数, レビュー時間とは負の相関があった.

この結果は, レビュー対象 B ではルールに関わらずゲーミフィケーションを好むほうが, 指摘数が増える傾向があり, ルール β の場合にレビュー時間が減る傾向があることを示している. 逆にレビュー対象 A の場合, ゲーミフィケーションを好む場合, レビュー時間, レビュー指摘数ともに減少する傾向があることを示している.

レビュー結果の正規化: レビュー能力には個人差があるため, 上記の分析では, 個人ごとにレビューの結果がルールによってどのように変化したのかが明確でない. そこで, レビュー対象 A の結果を分母, B の結果を分子としてレビュー時間, レビュー指摘数の変化(比率)を個人ごとに正規化して分析した. 例えば, レビュー対象 A での指摘数が 2, レビュー対象 B での指摘数が 4 の場合, 正規化の

結果は $4/2=2$ となる。

Q1 と正規化したレビュー結果との相関係数を表 3 に示す。ルールによりレビュー時間が変化すると仮定すると、ルール α をレビュー対象 A, ルール β をレビュー対象 B に適用した場合、前者のレビュー時間が長くなり、後者のレビュー時間が短くなる。ゲーミフィケーションを好むほどこの傾向が強い場合、正規化したレビュー時間は小さくなるため、相関係数は負の値となる。逆にルール α をレビュー対象 B, ルール β をレビュー対象 A に適用した場合、相関係数は正の値となる。

表において、どちらの場合も相関係数が正の値となったことから、ゲーミフィケーションを好む場合、規模の大きいレビュー対象 B に、より時間を掛けていることとなる。このことから、ルールによりレビュー時間が変化することはないといえる。レビュー指摘数についても同様で、どちらの相関係数も正の値となっていた。このことから、ルールに時間の要素が含まれていても、レビュー指摘数が少なくなる可能性は低いと考えられる。

レビュー効率の分析: さらにレビュー効率についても分析した。ここではレビュー効率は、レビュー指摘数 / レビュー時間と定義する。すなわち、時間あたりの指摘数が多いほど、効率が高いとする。レビュー対象 A にルール α , レビュー対象 B にルール β を適用する場合などに、ルールの違いによりレビュー効率が変化するのかを確かめた。

分析結果を表 4 に示す。レビュー対象 B にルール β を適用した場合、ルール α を適用した場合よりもレビュー効率が高くなっていた。逆にレビュー対象 A に関しては、ルール β を適用した場合にレビュー効率が低下しており、結果が一貫していなかった。このことから、ルールに時間の要素が含まれていても、レビュー効率が高まる効果はないと考えられる。

リサーチクエスチョンに対する答え: 上記の結果より、ゲーミフィケーションを適用することにより、レビュー指摘数を若干高める効果が期待できるといえる。実験で適用したルールでは、レビュー時間を短くする効果はなかった。ただし、レビュー時間が短くなり、その結果レビュー指摘数も減少するという悪影響も見られなかった。よって、RQ1-1 に対する答えは「増加する」、RQ1-2 に対する答えは「(適用したルールでは) 短縮されない」、RQ1-3 に対する答えは「(適用したルールでは) 悪影響を受けない」となる。

5.3 コーディング

コーディング時間とゲーミフィケーションのルール（質問項目 Q1～Q3）との関係、及び時間割引率との関係を分析した。分析では、それぞれのルールを重視するほどコーディング時間が変化すると仮定し、コーディング時間と各ルール（質問項目）との相関係数を算出した。コーディングタスクごとにルールの組み合わせを変えて実験しており、各ルールの効果を確かめるためにはデータを層別する必要

表 5 各項目とコーディング時間との相関係数

	Q1	Q2	Q3	時間割引率
全体	-0.37	0.00	0.09	0.12
仕様 A	-0.07	0.08	-0.09	-0.22
仕様 B	-0.44	-0.04	0.15	0.26

表 6 各項目とコーディング時間との相関係数

(a) 仕様 A にルール δ を適用				
	Q1	Q2	Q3	
仕様 A	0.10	0.33	0.28	-0.18
仕様 B	-0.62	0.03	-0.06	0.16
(b) 仕様 A にルール γ を適用				
	Q1	Q2	Q3	
仕様 A	-0.24	-0.09	-0.37	-0.34
仕様 B	-0.21	-0.09	0.48	0.39

表 7 各項目と正規化したコーディング時間との相関係数

ルール γ 適用対象	Q1	Q2	Q3	時間割引率
仕様 A	-0.58	-0.17	-0.31	0.34
仕様 B	0.13	0.04	0.76	0.34

がある。そこで前節と同様に、タスクとルールの組み合わせによりデータを層別して相関係数などを算出した。

ゲーミフィケーションの効果: まず、ゲーミフィケーション自体の効果について確かめた。コーディング時間と各項目との相関係数を表 5 に示す。表で「全体」とは層別せずに全被験者で相関係数を算出した場合、及び各仕様で層別した場合の相関係数を示す。質問項目 Q1 との関係を見ると、全て負の相関となっていた。仕様 A では相関係数が小さいが、これは仕様 A が簡単であったため、差がつきにくかった可能性がある。この結果より、ゲーミフィケーションはコーディング時間を減らせる可能性があるといえる。

ゲーミフィケーションのルールの影響: 次に、ゲーミフィケーションのルールの影響を分析するため、仕様とルールの組み合わせによりデータを層別した場合の結果を表 6 に示す。質問項目 Q1 とコーディング時間との関係では、仕様 A とルール δ の組み合わせを除き、相関係数が負の値となっており、かつ比較的大きな値となっていた。すなわち、ルールに関わらずゲーミフィケーションを好むほうが、コーディング時間が減少していた。Q2, 時間割引率については、ルールを入れ替えているにも関わらず相関係数の正負が変化していないことから、コーディング時間と関係がないといえる。質問項目 Q3 については、ルール δ (前回得点へ加算するルール) を適用している場合に、ルール δ を重視する被験者のコーディング時間が逆に長い傾向があった。

コーディング時間の正規化: コーディング能力の個人

差は大きいと考えられるため、仕様 A と B のコーディング時間の比に着目して分析を行った。具体的には、仕様 B のコーディング時間 / 仕様 A のコーディング時間を算出し、その値と各質問項目などとの相関係数を算出した。これにより、仕様 A を基準として、ルールの違いにより仕様 B のコーディング時間がどのように変化しているかを分析することができる。ここで仕様 A にルール δ 、仕様 B にルール γ を適用しており、ルール δ に効果がある場合、前述の比の分子の値（仕様 A のコーディング時間）が小さくなり、分母が大きくなることから、比の値も大きくなる。ゲーミフィケーションを好むほどこの傾向が強い場合、この比と各質問項目との相関係数は正の値となる。

結果を表 7 に示す。時間割引率については、ルールに関わらず相関係数が正の値となっていたことから、ルールとは関係がないといえる。その他については、前述のようにルール δ に効果があるならば、仕様 A にこのルールを適用した場合、各質問項目との相関係数は正の値となるが、実際には逆の傾向となっていた。すなわち、被験者の主観的評価とは逆の傾向が見られた。

リサーチクエスチョンに対する答え：実験結果より、ゲーミフィケーションによりコーディング時間を減少できる可能性があるといえる。スコアをリアルタイムで表示する、前回の得点に加算するなどのルールについては、実際に時間を短縮する効果は低いと考えられる。よって、RQ2-1 に対する答えは「短縮する」、RQ2-2 に対する答えは「影響に違いはない」となる。

5.4 ゲーミフィケーションの主観的評価

被験者がゲーミフィケーションを主観的にどのように評価しているかを分析した。実際のコーディング時間短縮の効果が小さくても、ゲーミフィケーションを主観的には評価している場合、作業に対するモチベーションの向上につながり、長期的には何らかの効果が生まれる可能性がある。

質問項目 Q1 から Q3 について、平均、標準偏差、及び回答が 2 以下となった割合を表 8 に示す。Q1、Q2 については、平均値が 3.5 を上回っており、2 以下の評価も多くないことから、比較的評価が高いといえる。Q3 については、43% が 2 以下の評価をしていることから、個人によってルールに対する評価差があるといえる。なお Q1 については、2 以下と評価された評価が 29% あり、ゲーミフィケーションを好まない被験者がある程度存在していることを示している。このため、分析や実際の開発への適用において、ゲーミフィケーションに対する個人の好みを考慮すべきであるといえる。

分析結果より、RQ3-1 に対する答えは「おおむね肯定的評価であるが、個人による差は無視できない」、RQ3-2 に対する答えは「ルールによっては個人ごとの評価が大きく異なる場合がある」、RQ3-3 に対する答えは「時間割引率と前

表 8 各質問項目の基本統計量

	Q1	Q2	Q3
平均値	3.5	3.7	3.1
標準偏差	1.3	1.0	1.3
2 以下の回答割合	29%	14%	43%

表 9 Q1 との相関係数

(a) コーディング時間		
仕様 A	仕様 B	全体平均
-0.03	-0.44	-0.35
(b) レビュー時間		
レビュー対象 A	レビュー対象 B	全体平均
-0.47	0.30	0.05

表 10 時間割引率と各質問項目との相関係数

Q1	Q2	Q3
-0.10	-0.19	-0.38

表 11 各開発作業時間の個人差

(a) コーディング時間		
仕様 A	仕様 B	
標準偏差	0:04	0:10
平均	0:10	0:32
(b) レビュー時間		
レビュー対象 A	レビュー対象 B	
標準偏差	0:03	0:07
平均	0:13	0:25

回の結果を合算するルールに関連が見られる」となる。

5.5 ゲーミフィケーションの影響要因

作業時間との関係：ゲーミフィケーションに影響する要因について分析した。まず、質問項目 Q1、すなわちゲーミフィケーションを好むことの要因に着目する。単純にレビュー時間やコーディング時間が短い、すなわち作業が早い被験者がゲーミフィケーションを好む可能性がある。そこで、Q1 と各時間との相関係数を算出した。結果を表 9 に示す。ここで全体平均とは、例えば仕様 A と仕様 B のコーディング時間の平均を指す。コーディング時間の場合、規模の大きい仕様 B、全体平均について相関係数が負の値となっていた。逆にレビュー時間の場合、レビュー対象 B、全体平均の相関係数が正の値となっていた。このことから、単純に作業が早い被験者がゲーミフィケーションを好むわけではないといえる。

前回の結果との関係：次に、前回のゲーミフィケーションの結果、すなわち、レビューにおけるゲーミフィケーションの順位（結果）が良かった場合に、ゲーミフィケーションを好むのかどうかを分析した。具体的には質問項目 Q1、Q3 と前回の順位との関連を確かめた。その結果、それぞれの相関係数は 0.17、0.05 となり、0.2 を下回っていた。このことから、前回の結果によってゲーミフィケーションそ

のものやゲーミフィケーションのルールに対する好ましさが変化しているわけではないといえる。

時間割引率との関係：最後に時間割引率とゲーミフィケーションとの関係について分析した。時間割引率は結果を早く求める傾向があるとも解釈できるため、ゲーミフィケーションと何らかの関係を持つ可能性がある。時間割引率と質問項目 Q1～Q3 の相関係数を表 10 に示す。Q1, Q2 との相関係数は小さかったが、Q3 との相関係数は 0.3 を超えており、かつ正の値となっていた。このことから、時間割引率が高い場合、前回の得点に今回の得点を加算するルールを好む傾向があるといえる。

5.6 コーディングとレビューの関係

コーディング時間とレビュー時間などに関係があるかどうか、また、各時間の個人差はどの程度かを分析した。コーディング時間、レビュー時間の平均と標準偏差を表 11 に示す。コーディング、レビューとも時間の平均が同程度の場合、標準偏差も同程度であった。このことから、コーディングとレビューの個人差のばらつきに、大きな差はないと考えられる。

次に、コーディング能力とレビュー能力との関係を分析した。コーディング時間とレビュー時間の相関係数は -0.21、コーディング時間とレビュー指摘数の相関係数は 0.01 であった。相関係数は小さく、また、レビュー時間とコーディング時間との相関係数は負の値となっていたことから、コーディングが早い被験者でも、レビューが早く、指摘数も多いとは限らないといえる。

6. おわりに

本研究では、ゲーミフィケーションのルールの差異が作業の成果に影響するかどうか、また、ゲーミフィケーションのルールに対する評価が個人により異なるかどうかを分析した。分析から、新たに以下の知見が得られた。

- レビュー時間短縮を狙ったゲーミフィケーションのルールを適用した場合、時間短縮の効果は見られなかったが、欠陥見逃しが増えるといった傾向も見られなかつた。
- ゲーミフィケーションのルールの差異は、主観的評価に対しては影響を与えていたといえる。また、個人によりルールに対する評価が異なるため、個人差を考慮してルールを設定する必要がある。
- 時間割引率が高い場合、前回の得点に今回の得点を加算するルールを好む傾向が見られた。

今後の予定は、ルールをさらに変化させ、ソフトウェア開発の効率や品質に影響するかどうかを分析することである。

謝辞 本研究の一部は、文部科学省科学研究補助費（基盤 C：課題番号 16K00113）による助成を受けた。

参考文献

- [1] AIZU ONLINE JUDGE, <http://judge.u-aizu.ac.jp/>
- [2] Dubois, D., and Tamburrelli, G.: Understanding gamification mechanisms for software development, In Proc. of Joint Meeting on Foundations of Software Engineering (ESEC/FSE 2013), pp. 659-662 (2013).
- [3] 薗間文彦: アンケートによる時間割引率の背景要因に関する研究, 早稲田商学, vol.432, p.1-34 (2012).
- [4] 一ノ瀬智浩, 上野秀剛: ゲーミフィケーションを構成する要素の違いと作業効率の評価, ヒューマンインタフェース学会論文誌, vol.18, no.2, pp.65-76 (2016).
- [5] Khandelwal, S., Sripada, S., and Reddy, Y.: Impact of Gamification on Code review process: An Experimental Study, In Proc. of Innovations in Software Engineering Conference (ISEC '17), pp.122-126 (2017).
- [6] 根本紀之: ゲーミフィケーションを用いた探索的テストの効果報告, ソフトウェア・シンポジウム 2016 in 米子, pp.47-52 (2016).
- [7] 應治沙織, 上野秀剛: コードレビュー時の読み方教示によるレビュー効率の向上, 情報処理学会研究報告 ソフトウェア工学研究会, vol.2014-SE-185, no.2, pp.1-8 (2014).
- [8] Smith R., and Kilty, L.: Crowdsourcing and Gamification of Enterprise Meeting Software Quality, In Proc. of International Conference on Utility and Cloud Computing (UCC '14), pp.611-613 (2014).
- [9] Stanculescu, L., Bozzon, A., Sips, R., and Houben, G.: Work and Play: An Experiment in Enterprise Gamification, In Proc. of ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16), pp.346-358 (2016).
- [10] Unkelos-Shpigel, N., and Hadar, I.: Gamifying software engineering tasks based on cognitive principles: the case of code review, In Proc. of International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE '15), pp.119-120 (2015).