

年齢がコードレビューに与える影響の分析

Influence of the Reviewers' Age to Software Code Review

村上 優佳紗† 角田 雅照† 上野 秀剛‡
Yukasa Murakami Masateru Tsunoda Hidetake Uwano

1. はじめに

ソフトウェア開発者の能力は個人により大きく異なることが、いくつかの研究で指摘されている。例えば、Sackman らの研究において、プログラムのデバッグ能力が、個人により大きな差があることが示されている[5]。同様に、Thelin らの研究においても、個人によりコードレビュー能力の差が大きいことが示されている[6]。もし開発者の能力差を小さくする、すなわち、能力が比較的低い開発者について、ある程度能力を高めることができれば、ソフトウェア開発プロジェクトの生産性を改善できる可能性がある。

開発者の能力差を小さくするためには、能力差に影響する要因を明らかにし、その要因の影響を小さくするためのソフトウェア工学教育やソフトウェア開発支援ツールを用意する必要がある。例えば、個人によりコードレビュー能力に差があり、その原因が仕様と実装コードの対応をチェックしていないことが原因である[1]とする。この場合、「仕様と実装コードの対応をチェックする」ことをレビューのチェックリストに追加することにより、レビューの能力差を小さくすることができる。これまで、開発者の能力差の要因について分析した研究がいくつか存在する[2]。

本研究では、開発者の能力差の要因として、開発者の年齢に着目する。科学的な根拠は希薄であるが、開発者の能力と年齢の関連が指摘される場合がいくつか見られる。例えば、ある記事[8]では、年齢が若い開発者のほうが、新しいコーディング方法や技術の習得が早いと指摘している。また日本では、35歳でプログラマとしての能力の限界を迎えるという「35歳限界説」[9]が散見される。ただし、実際に年齢により開発者の能力が影響されるのかどうか、影響されるとすればどの能力なのかは、我々の知る限り、これまでほとんど分析されていない。

もし年齢と開発者の何らかの能力に関連がある場合、開発者の年齢に則して開発ツールの支援方法、支援内容を変更する必要がある。そのようなツールを提供できれば、年齢の高い開発者とそうでない開発者の能力差を小さくすることができ、それによりソフトウェア開発プロジェクトの生産性を高められる可能性がある。

†近畿大学, Kindai University

‡奈良工業高等専門学校,
National Institute of Technology, Nara College

開発者の取り組む作業として、本研究ではソースコードレビューに着目して分析する。すなわち、ソースコードレビューの効率、正確性などにおいて、被験者の年齢が何らかの影響を及ぼすのかを分析する。一般的には、年齢が高いほうがコーディングの経験が多くなるため、ソースコードレビューに良い影響を与えると考える。ただし、我々のこれまでの研究[3]では、ソースコードの理解速度については、記憶力を必要とするものについては、被験者の年齢が高い場合、比較的遅い傾向が見られた。そのため、コードレビューにおいても、効率（時間あたりの指摘数）などが年齢に影響を受ける可能性がある。

2. 関連研究

ソフトウェア開発者の能力は個人により大きく異なることが、いくつかの研究で指摘されている[2][5]。例えば、Sackman らの研究において、プログラムのデバッグ能力が、個人により大きな差があることが示されている[5]。ただし、年齢が開発者の能力に影響するのかどうか、影響するとすればどの能力なのかについて分析した研究は、我々の知る限り、Morrison ら[2]の研究など、ごく少数である。

Morrison らはソフトウェア開発者の知識と年齢が関係するかどうかを Stack Overflow という Q&A サイトのデータを用いて分析している。その結果、開発者の評価が年齢とともに上昇することなどを示している。ただし、Morrison らの研究では開発者の知識に着目しているが、記憶力には着目しておらず、年齢とプログラムの理解速度の関係などは明らかにしていない。

3. 準備

3.1. レビュー対象コード

レビュー対象のソースコードとして、練習問題と本番問題の 2 種類を用意した。練習問題は実験で用いるツールの操作に慣れることを目的としている。練習問題は以下の構成とした。

- 設計書 (約 100 字)
- Main.java (約 10 行)
- Sum.java (約 10 行)

本番問題は、練習問題よりも分量を増やした以下の構成とした。

- 設計書 (約 600 字)
- Main.java (約 100 行)

表 1 被験者の基本データ

		年齢	経験年数	コーディング時間	最大規模	Java 経験年数	Java レベル
高年齢	平均値	42.4	26.4	10.0	24833.3	8.9	2.9
	中央値	40.5	27.5	6.0	10000.0	8.0	2.5
低年齢	平均値	21.3	4.3	16.2	1337.3	3.2	3.0
	中央値	21.0	3.0	20.0	1012.0	3.0	3.0

- House.java (約 30 行)
- bukken.dat (約 20 行. プログラムで読み込むデータ)

被験者は説明文を読んでからレビューに取り組む。被験者は設計書を読み、ソースコードの修正箇所をレビュー指摘書に記入する。修正箇所の数は被験者に伝えず、制限時間を設けていない。そのため、被験者本人が修正する箇所がないと判断した時点で実験終了とする。

レビュー時間などを計測するため、各シートの表示回数と表示時間を記録するツールを作成した。ツールは表計算ソフトのマクロを用いて開発した。以下の複数のシートから構成される。

- 被験者の年齢などのアンケート
- 実験概要の説明
- 設計書
- ソースコード (各クラスに 1 シート)
- レビュー指摘書

被験者は適宜シートの表示を自分で切り替える。本番問題のソースコードなどは、練習問題の完了後に、練習問題と入れ替わりで表示される。

3.2.分析対象のメトリクス

以下のメトリクスを用いて、レビューの効率及び正確性に影響する要因を分析した。時間に関するメトリクスの計測単位は「分」である。

- 年齢: 被験者の年齢
- 経験年数: 被験者のプログラミング経験年数
- コーディング時間: 1 週間でコードを読む、または書く時間の平均。単位は時間
- 最大規模: 被験者がこれまでに開発したソフトウェアの最大規模
- Java 経験年数: 被験者の Java の経験年数
- Java レベル: レベル 1~4. 数値が大きいほど Java に熟達していることを示す。自己申告に基づく
- アンケート 1: 実験後アンケート。最初に設計書をよく読んだかどうか。数値が小さいほど当てはまっていることを示す。アンケート 2, 3 も同様
- アンケート 2: コード全体にざっと目を通したかどうか
- アンケート 3: 設計書に書かれているフィールド、メソッドが正しくソースコードに実装さ

れているか確認したかどうか

- レビュー時間: レビュー完了までに掛かった時間
- 正答数: 問題のある箇所を指摘した合計
- 指摘数: レビューの間に指摘した箇所の合計
- 正答率: 正答数 ÷ 問題のある箇所の合計
- 正答含有率: 正答数 ÷ 指摘数
- レビュー指摘時間: 指摘数 ÷ レビュー時間
- 正答指摘時間: 正答数 ÷ レビュー時間

分析では、正答率、正答含有率、レビュー指摘時間、正答指摘時間を目的変数とし、その他の変数は主に説明変数として扱った。

3.3.被験者

年齢に着目して分析するため、被験者を低年齢グループと高年齢グループに分けた。低年齢グループは、近畿大学理工学部情報学科に所属する学部 4 年生及び修士 1 年生の 6 名である。大学の実習科目において設計書を読み実装を行うことや、設計書を作ることを経験しているため、それらの作業にある程度習熟している。全員が設計書のある程度読んだことがある、また書くことができる。高年齢グループは、同学科の教員 7 名と社会人 1 名の計 8 名であり、うち教員 2 人はソフトウェア開発企業での実務経験を持つ。

被験者の基本データを表 1 に示す。高年齢グループのほうが経験年数、最大規模、Java 経験年数とも高くなっている。ただし、コーディング時間と Java レベルは低年齢グループのほうが高くなっていた。

4. 分析

低年齢グループと高年齢グループでデータを層別した場合の、レビューに関するメトリクスを表 2 に示す。回答時間の平均値、中央値とも低年齢グループと高年齢グループでほとんど差がなかったことから、どちらのグループもある程度レビューに丁寧に取り組んだと考えられる。低年齢グループの正答指摘時間の平均値がやや高かったが、中央値は低かった。レビュー指摘時間にもほとんど差がなかったことから、レビュー効率は高年齢グループと低年齢グループではおおむね同じであるといえる。コードリーディングに関する従来研究[3]では、高齢者グループはコードの理解に記憶力を必要とする場合、速度が遅くなる傾向が見られたが、レビューではそのよ

表2 年齢により層別した場合の各メトリクス

		レビュー時間	正答数	指摘数	正答含有率	正答率	レビュー指摘時間	正答指摘時間
高年齢	平均値	28.4	3.1	9.4	36.1%	44.6%	3.2	6.6
	中央値	25.6	3.0	9.0	33.2%	42.9%	3.1	7.0
低年齢	平均値	27.5	3.3	8.0	38.3%	47.6%	3.3	8.3
	中央値	21.5	3.0	8.0	41.7%	42.9%	3.5	5.9

うな傾向は見られなかった。

レビューの正確性について、正答数を見ると高年齢グループと低年齢グループでほとんど差がなかった。正答率にもほとんど差がなく、正答含有率は低年齢グループのほうが若干高かった。高年齢グループは、表1に示すように経験年数、最大規模、Java経験年数とも高かったが、これらはレビューの正確性に強く関連していない可能性がある。

4.1. 考察

実験で用いたプログラムに含まれる誤りをその特性により分類し、高年齢グループと低年齢グループで発見する誤りに違いがあるかどうかを分析した。誤りの特性は以下のように分類した。

- 完全性: 設計された内容は完全にコーディングされているか
- 初期化: 変数とパラメータが初期化されているかなど
- 演算: 不等号の向きや等号の有無は正しいかなど
- データ・ファイル: データの型、値は正しく設定されているかなど
- 制御: ループや if 文は正しく開始、終了するかなど

誤りの特性と各グループの正答率を表3に示す。誤りの特性をみると、完全性と制御では高年齢グループの正答率が高かった。初学者グループでは、演算(2つあった誤りのうちの1つ)や初期化の正答率が比較的高かった。データが少ないため、さらに分析を行う必要があるが、高年齢グループでは経験が多い被験者が多かったため、経験が働いて見つけることができる誤りがあった一方で、経験による思い込みで見逃してしまった誤りが存在した可能性がある。

5. おわりに

本研究では、ソフトウェア開発者の能力に影響する要因のひとつとして、開発者の年齢に着目し、ソフトウェアレビューにおいて、年齢がレビューの正

確性や効率に影響するかどうかを分析した。その結果、被験者の2つのグループ、すなわち20歳代と40歳代において、年齢の高低はレビューの正確性や効率に影響していなかった。

今後の予定は、被験者を増やし結果の信頼性を高めるとともに、レビュー以外に年齢に影響されるソフトウェア開発活動を特定し、年齢の影響の大きさを明らかにすることである。

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。

6. 参考文献

- [1] 栗山進, 大平雅雄, 門田暁人, 松本健一: プログラム理解度がコードレビュー達成度に及ぼす影響の分析, 電子情報通信学会技術報告, SS2004-53, pp.17-22 (2005).
- [2] Morrison, P. and Murphy-Hill, e.: Is programming knowledge related to age? an exploration of stack overflow, In Proc. of Working Conference on Mining Software Repositories (MSR), pp.69-72 (2013).
- [3] 村上優佳紗, 角田雅照, 中村匡秀: ソフトウェア開発者の年齢がプログラム理解速度に及ぼす影響の分析, 情報処理学会研究報告, 2016-SE-191, No.1, p.1-6 (2016).
- [4] 應治沙織, 上野秀剛: コードレビュー時の読み方指示によるレビュー効率の向上, 情報処理学会研究報告, 2014-SE-185, No.2, pp.1-8 (2014).
- [5] Sackman, H., Erikson, W. and Grant, E.: Exploratory experimental studies comparing online and offline programming performance, Communications of the ACM, vol.11, no.1 (1968).
- [6] Thelin, T., Andersson, C., Runeson, P., Dzamashvili-Fogelström, N.: A Replicated Experiment of Usage-Based and Checklist-Based Reading, Proc of International Symposium on Software Metrics, pp.246-256 (2004).
- [7] Uwano, H., Nakamura, M. Monden, A. and Matsumoto, K.: Exploiting Eye Movements for Evaluating Reviewer's Performance in Software Review, IEICE Transactions on Fundamentals,

表3 誤りの特性と正答率との関係

	完全性	データ・ファイル	演算 1	演算 2	初期化	制御
低年齢	33%	17%	50%	83%	83%	17%
高年齢	75%	13%	25%	88%	63%	38%

vol.E90-A, no.10, pp.317-328 (2007).

- [8] Wadhwa, V.: Silicon Valley's Dark Secret: It's All About Age, <http://techcrunch.com/2010/08/28/silicon-valley%E2%80%99s-dark-secret-it%E2%80%99s-all-about-age/> (2010).
- [9] 山田モーキン: Web エンジニア「35 歳限界説」は本当か? http://next.rikunabi.com/01/closeup_1251/index.html

1