

## HDR 法：仮説駆動型レビュー手法の提案

- HDR 法の実践による生産性と品質の同時向上 -

### Proposal of HDR method. (HDR: Hypothesis Driven Review)

The best practice to adopt the HDR method to improve both productivity and quality.

2012 年度 SQIP 研究会第 3 分科会

The 3rd Section Meeting Team of SQIP Study Group in 2012

主査	細川 宣啓 <sup>1)</sup>			
副主査	永田 敦 <sup>2)</sup>	藤原 雅明 <sup>3)</sup>		
アドバイザー	森崎 修司 <sup>4)</sup>			
研究員	○高橋 功 <sup>5)</sup>	上田 裕之 <sup>6)</sup>	高橋 実雄 <sup>7)</sup>	中谷 一樹 <sup>8)</sup>
CHIEF INVESTIGATOR	: Nobuhiro Hosokawa <sup>1)</sup>			
SUB-CHIEF INVESTIGATOR	: Atsushi Nagata <sup>2)</sup>	Masaaki Fujiwara <sup>3)</sup>		
ADVISER	: Shuji Morisaki <sup>4)</sup>			
RESEARCHER	: ○Isao Takahashi <sup>5)</sup>	Hiroyuki Ueda <sup>6)</sup>		
	Mitsuo Takahashi <sup>7)</sup>	Kazuki Nakatani <sup>8)</sup>		

#### 研究概要

本研究チームでは、レビューアの思考メカニズムに着目し、時間制約の中でも集中的に重大欠陥を検出できるレビューア（以降、トップレビューア）の思考メカニズム解明を研究対象とした。

レビューアの思考メカニズムを解明した結果、トップレビューアはレビュー対象および品質要求・品質特性から欠陥の兆候を掴み、兆候を基に仮説を立てて狙いを定めてから検証を行っていることが判明した。本研究では、トップレビューアが普段から行う暗黙知としての思考メカニズムを、手法として形式知化し初心者／未経験者でも実践できるように体系化した。

本手法を用いた実験により、効率的で効果的なレビューの実施、レビューアのスキル向上を図ることが可能になるという確信を得た。

#### Abstract

Our research team focused on reviewers' thinking mechanism and above all, researched the top reviewers' which helped them detect critical defects intensively even with a time restriction.

By clarifying the thinking mechanism, we found that they grasp a sign of defects from review objectives or quality requirement and quality characteristics, then focus and verify it after building up a hypothesis based on

---

ソーバル株式会社 品質評価部

Sobal Corporation, Verification&Validation Div.

東京都大田区下丸子 3-25-14 Tel:03-5482-1222 e-mail:isao\_takahashi@sobal.co.jp

3-25-14, Shimomaruko, Ota-ku, Tokyo Japan Tel:03-5482-1222

1) 日本アイ・ビー・エム株式会社

IBM Japan, Ltd

2) ソニー株式会社

SONY Co, Ltd

3) 東芝ソリューション株式会社

TOSHIBA Solution Co, Ltd

4) 国立大学法人静岡大学

National University Corporation Shizuoka University

5) ソーバル株式会社

Sobal Corporation

6) 株式会社 DTS

DTS Corporation

7) サントリーシステムテクノロジー株式会社

Suntory System Technology, Ltd

8) TIS 株式会社

TIS Inc

the sign of defects.

In this research, we formalized the tacit about quality, that is, thinking mechanism that they usually apply as a method, and then defined the process to enable beginners or inexperienced people to practice it.

The experiment using this method gained us confidence that it enabled to carry out an effective and efficient review and to improve reviewers' skill.

## 1. はじめに

### 1.1 研究の背景とレビューに関する問題点

近年のソフトウェア開発において、レビューはソフトウェアの欠陥を早い段階で検出できる手段として、品質向上に寄与するだけでなく、コスト削減、納期短縮に有用である。

しかし、レビューは属人性が高い。つまり、手法の自由度が高く、その効果が個人のレビュースキルに依存する傾向が強いのである。よって重要なレビューの依頼は、レビュースキルの高いレビューアに集中してしまう傾向にある。そのような負荷集中を解消するために、他の要員にレビューを分担するが、欠陥検出能力が低いため多くの時間を費やしても軽微欠陥しか検出できないことが多い。

また、近年のソフトウェアは規模が大きくなり複雑化している。ソフトウェアの複雑度が増すことでレビューで確認すべきポイントは指数関数的に増加し、限られた時間の中で全てのレビュー対象を全ての確認ポイントでレビューするという事は不可能になりつつある。開発期間も短縮傾向にある中、レビューに費やすことができる時間が限られてきている。

これらが原因となり、重大欠陥が残存し、レビューしているにも関わらず手戻りが発生して納期遅延や工数増大を招いたり、本番障害が発生してシステム利用者からの信頼喪失や重大事故に発展したりしてしまう。我々はこの「重大欠陥の残存」を最も重要な問題と捉え、その解決方法を検討した。

### 1.2 研究の狙い

重大欠陥残存の問題へのアプローチとして、レビュープロセスやレビュー観点の改善などが考えられるが、我々は「個人の欠陥検出能力向上」に着目した。

誰でも時間制約がある中で集中的に重大欠陥を検出できるようになること、更にはその技術を移転可能な形にし、欠陥検出能力の向上が容易となることで問題を解決できると考えた。

そこで着目したのが、様々なプロジェクトにおいてレビューによる大きな成果を上げているレビューアのレビュースキルである。

時間制約がある中でも重大欠陥を検出できる、いわゆる「トップレビューア」の欠陥検出メカニズムを研究の対象とし、トップレビューアがレビュー時に何を見て何を考えているのか、その思考メカニズムを解明し形式化して新たなレビュー方法、「HDR法 (Hypothesis Driven Review Method: 仮説駆動型レビュー手法)」を確立した。

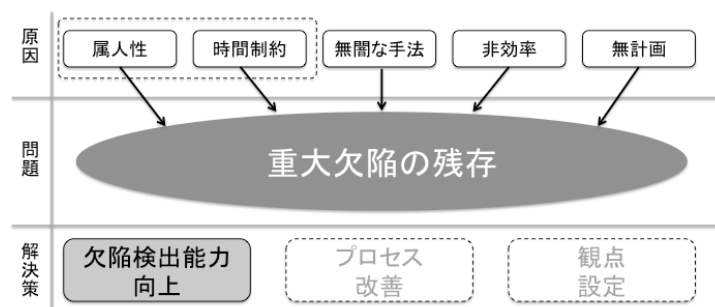


図 1-1 レビューで抱える問題点と解決策

## 2. 提案する欠陥検出の方法

本章では、我々の研究グループが提唱する「HDR法」について記述する。本手法を構成する要素である「思考プロセス」と、本手法を使いこなすための「必要な知識」について説明する。

通常、レビューを行う際は1ページ目の1行目から順番にレビュー対象を読み込んで行くことが殆どである。しかし、トップレビューはこれとは全く異なるアプローチを取っていることが分かった。レビュー対象の概観やその周辺を俯瞰することから始め、全体を流し読みする。そして、レビュー対象の特徴から欠陥の兆候を掴み、兆候を基に混入欠陥の仮説を立てて狙いを定めてから検証（欠陥検出）を行っていた。

その思考プロセスは以下（図2-1）のようになる。

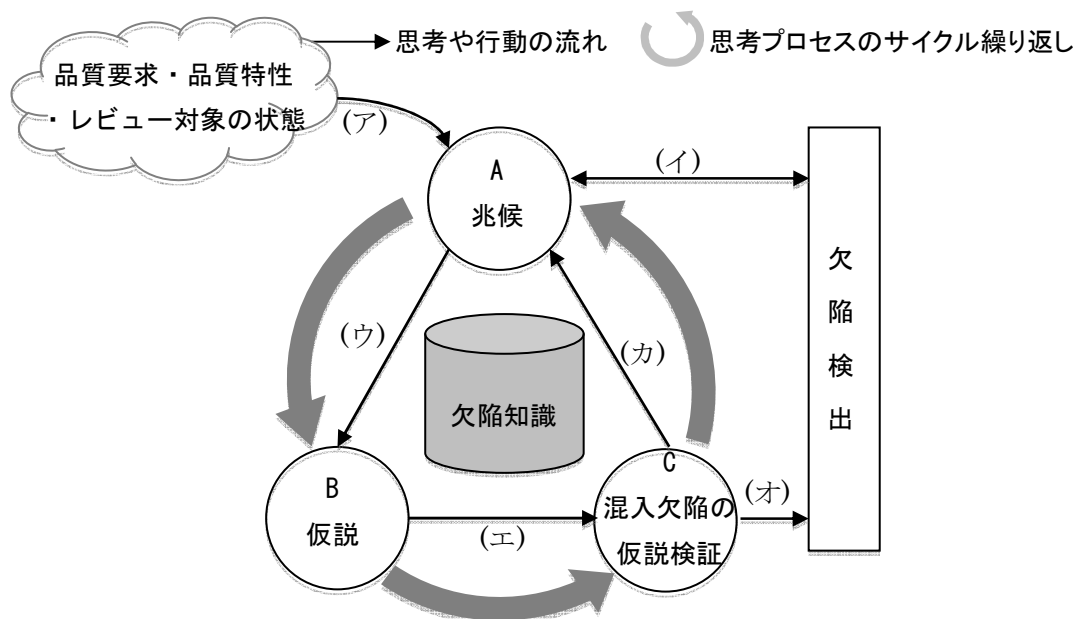


図 2-1 HDR 法概念図

以下に思考プロセスの詳細を示す。

### (1) 兆候を掴む

レビュー時の第一の行動として、レビュー対象の中身を詳細に見るのではなく、兆候を掴むことから開始する。兆候とは、欠陥の存在を示唆する定量的・定性的なデータやキーワードなど、レビュー対象の内外に観察可能な「特徴」のことである。

例えば、ドキュメント内に「”、や。（句読点）”と”、や。（カンマ、ドット）”が混在している」、または「である調，ですます調の混在」、「文字間，行間が一部だけ詰まっている。フォントが違う」、「インデント幅が一部だけ異なる」というようにレビュー対象中の表現や書式が一部異なる特徴は、作成者以外の人間が後から追加した箇所があることを示唆する兆候となる。また、レビュー対象の作成時間（ファイルプロパティの作成日時から最終更新日時までの間隔）が極端に短いという特徴は、レビュー対象を急いで書き上げ、自己レビュー不足の状態であることを示唆する兆候となる。後者はレビュー対象の外側から掴むことができる兆候である。

ソースコードにも兆候は存在する。例えば「int i」というような変数名を使用している場合は、作成者のプログラミング経験が浅いということを示唆する兆候となる。

或は、納期要求や動作環境制約などの品質要求・品質特性（後述）をインプットにして兆候を掴む

こともできる(ア)。兆候と紐付けられる欠陥は一つではなく、N対Nの関係となる。

なお、明らかな欠陥として表出している兆候は直接欠陥検出に結びつく(イ)。この場合、その欠陥自体が他の欠陥の兆候となりうる可能性がある。

## (2) 仮説を立てる

(1) で掴んだ兆候から、混入欠陥の種類や位置、及び混入状態の仮説を一つ以上立てる(ウ)。仮説への根拠となる兆候も一つではなく複数の兆候を基にすることが多い。

通常、故意に欠陥を混入させる人はおらず、無意識のうちに欠陥は混入する。そこには人的ミスが発生しやすい状況など、何か原因がある。すなわち原因と結果の因果関係がある。

仮説を立てる際は、兆候を結果と捉えて、その結果が何故起こったのか、原因を推測する。または、兆候を原因と捉えて、その原因により何が起こるのか結果を推測する。

(1) の例の場合、「句読点とカンマ、ドットの混在」という兆候を結果と捉えると「作成者以外の人間が後から追記」が原因の一つとして推測できる。更にその推測を原因として捉えると、その前後で不整合が発生しているという仮説を立てることができる。

## (3) 混入欠陥の仮説検証

(2) で立てた仮説を検証するための観点を設定し(エ)、その観点到絞って実際にレビュー対象を検査する。その結果、欠陥の発見や(オ)、新たな兆候の発見(カ)につながる。

レビュー中は、以上の(1)～(3)の思考プロセスのサイクルを回し続ける。

なお、本手法はレビュー対象や見るべき観点を絞るプロセスであるとも言えるため、レビュー対象物の数や量が多くどこから手をつけて良いかわからないような場合、より高い効果を発揮する。ただし、観点を絞るため、軽微欠陥も含めるとレビュー結果における「指摘件数」は減る可能性がある。

## 2.2 HDR法で必要な知識

### (1) 欠陥知識

本手法において、欠陥知識は重要な要素の一つである。2.1章で述べた思考プロセスに沿ってレビューを実施してもトップレビューアと同じ速度・精度でレビューしていくことは難しい。トップレビューアは、豊富な欠陥知識を有しており、その知識を瞬時に取り出すことが出来るからである。

ただし、指摘一覧表に記載された欠陥の事象そのものを記憶しておくだけでは欠陥の兆候を見つけることはできない。抽象化された欠陥の事象と共に、それを引き起こす兆候についてもセットで記憶しておくことが重要である。

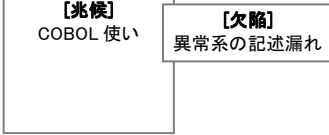
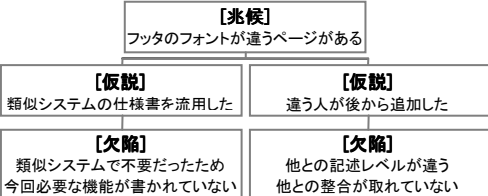
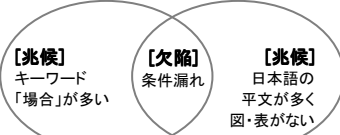
このような「整理・記憶された欠陥知識」と似た例として認知心理学のスキーマ[1]がある。人間は物事を認知する際、「特定の物体や場所の特性に関する知識」を利用することで、「目の前にある物体は何なのか」ということを理解する。この知識がスキーマである。

同じように、トップレビューアは欠陥知識を豊富に持っているからこそ、兆候を認知する能力に長けている。更に、その欠陥が「どのような欠陥か？」ということ「知覚」するための知識も兼ね備えているため、兆候からいくつもの仮説を導き出すことができるのである。

欠陥知識の蓄積には、日々のレビューにおいて、兆候-仮説-検証の思考サイクルを回すと共に、「どのような欠陥をどのように見つけたか？」を分析するための「振り返り」を欠かさず行い、日頃から意識的に整理・記憶していくことが重要である。そのためには、レビュー時に欠陥の事象だけではな

く兆候も同時に記録していくことや、表 2-1 に紹介するような欠陥知識の整理・記憶方法を実行することも有効である。

表 2-1 欠陥知識の記憶方法例

タグ型	階層構造型	ベン図型
<p>この兆候には、この欠陥と いうように、兆候に対してタ グを付ける。</p> 	<p>ツリー図などで、兆候→仮説→欠陥を図化 する。</p> 	<p>複数の兆候から一つの欠陥 の仮説を立てられるようなも のをベン図にする。</p> 

## (2) 品質要求・品質特性

レビュー対象物を作成する上での構築環境や前提条件・制約条件といった品質要求や品質特性は、兆候の一つとして、または仮説を立てる上での直接的な根拠となり、仮説立案のための重要な情報となる。間接的メトリクス[2]を用いて、プロジェクトや成果物の状態を計測することで、兆候を察知することも有効な手段の一つである。

## 3. 評価と結果

### 3.1 実験方法

実験では従来の自由なレビュー方法と、HDR 法でレビューした場合の結果を比較し、HDR 法の有用性を評価した。レビュー対象として二つの要求仕様書、仕様書 X「アンケート記入システム」、仕様書 Y「ヘリコプター予約管理システム」をそれぞれ 5 ページ、約 2000 字の分量で、難易度および欠陥混入度合いが同程度となるようにして自作した。研究員が所属する会社の中から所属部門、経験年数など、それぞれ立場の異なる被験者 36 名を選抜し、レビューを 2 回実施した。時間制約のある状況下を想定し、レビュー時間は指摘記述も含めて 25 分間とした。1 回目は通常のレビュー方法で実施した。その後、HDR 法に関する説明を行った後、2 回目を実施した。本手法には欠陥知識を要するため、説明の際は HDR 法による欠陥検出の実例を説明することで被験者に欠陥知識をインプットすることとした。なお、被験者のスキル、およびレビュー対象の違いによる誤差を極力排除するため、被験者を 2 グループに分けて仕様書を入れ換えて実施した。(表 3-1 参照)

表 3-1 実験方法

	A グループ(15 名)	B グループ(21 名)	レビュー手法	レビュー時間
1 回目	仕様書 X	仕様書 Y	自由	25 分(指摘記述時間含む)
～ HDR 法に関する説明と欠陥知識の例示 ～				
2 回目	仕様書 Y	仕様書 X	HDR 法	25 分(指摘記述時間含む)

レビュー終了後、被験者には定量評価 5 項目、定性評価 3 項目からなるアンケート(図 3-2)に答えてもらい、実験終了後 1 回目と 2 回目の欠陥指摘リストと共に被験者より回収した。

### 3.2 評価方法

欠陥指摘リストから、3 種類の重大欠陥「手戻り工数大の欠陥(例えば、登録データ閲覧画面の追加が必要など、手戻り工数が 8 時間以上になるもの)」、「ビジネスニーズを満たせない欠陥(例えば、

へリポートが1つしかないものの、2台以上着陸できてしまうなど、業務上支障が出るもの)、 「テストで検出できない欠陥 (例えば、登録取消機能の欠落など、仕様書通りにテストするだけでは検出が難しいもの)」を抽出し、1回目と2回目における重大欠陥指摘数の推移を見る。アンケート結果からは、効果の度合いと改善点を得る。以上を総合的に評価し、本手法の効果を測定した。

### 3.3 実験結果

被験者 36 名から回収した欠陥指摘リストより、重大欠陥の指摘を抽出した結果を以下に示す。

図 3-1 の通り、「手戻り工数大の欠陥」、「ビジネスニーズを満たせない欠陥」、「テストで検出できない欠陥」の指摘数について、それぞれ 2 回目で増加しており、ノンパラメトリック検定の WilCoxon/Kruskal-Wallis の検定 (有意水準 0.1) において統計的有意差が得られた。HDR 法が重大欠陥検出に有効であり、手法としての再現性が認められたことを示唆する結果となった。

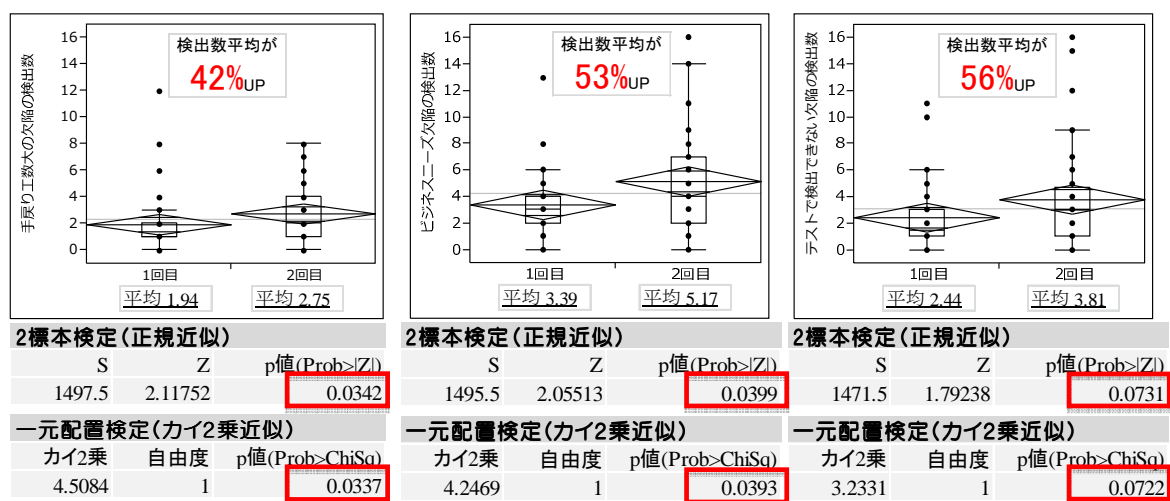


図 3-1 重大欠陥指摘数比較および統計分析結果

### 3.4 アンケート結果

被験者から回収したアンケート結果は以下の通りとなった。

#### (1) 定量評価結果

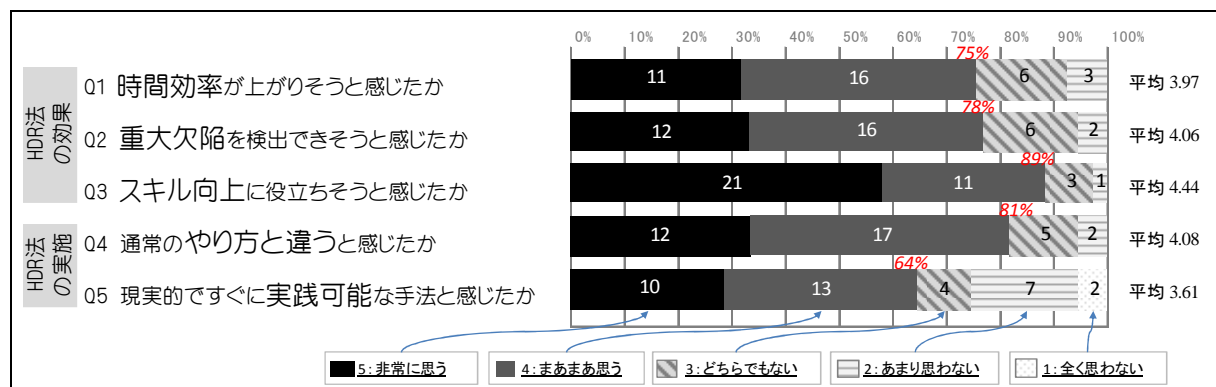


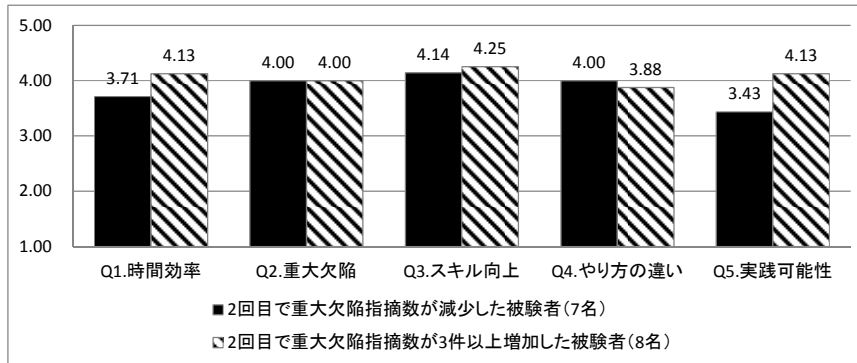
図 3-2 定量項目アンケート結果

設問 1~4 において、評価 4 以上が 75%以上という結果となり、特に「スキル向上に役立つ」と感じた被験者は 89%にも上った。設問 5 の「実践可能性」では評価 4 以上は 64%と比較的低い結果となった。

1 回目と 2 回目で「重大欠陥指摘数が 3 件以上増加した被験者」(8 名)と「重大欠陥指摘数が減少し

た被験者」(7名)がいる。これらの「著しく効果が出た被験者」と「効果の出なかった被験者」のアンケート結果を比較した(表3-2)。「時間効率」と「実践可能性」に対して、効果の出なかった被験者は低い評価となった。「やり方の違い」は効果の出なかった被験者の方がより違いを感じている。また、「重大欠陥」、「スキル向上」に対しては、両者共に比較的高い評価が得られた。

表3-2 効果が出た被験者と出なかった被験者の定量項目アンケート結果の平均



## (2) 定性評価結果

定性評価では、「兆候に着目することで全体を俯瞰することができるようになる」、「重大な欠陥を効率よく見つけることができる手法である」、「なぜ欠陥が混入したのか、その原因によって他にどんな欠陥が混入するのかを想起することは、レビュースキル向上に非常に役立つ」といった良い評価が多く挙げられた。

一方、「兆候から仮説を導く部分が難しく、知識や慣れが必要」という課題も挙げられた。特に前述の効果の出なかった被験者にこの意見が寄せられる傾向が見られた。

## 4. 考察

本章では第1章で挙げた「重大欠陥の残存」という問題に内在する「時間制約」、「属人性」に対する解決度合いについて、3章の実験結果を基に考察する。

### 4.1 時間制約に関する考察

既に3.3章で述べた通り、時間制約のある状況下を想定した実験において、自由手法と比較してHDR法を用いたレビューの方が、重大欠陥の指摘数は向上し、統計的有意差も得られた。

更にアンケート結果では定量評価で時間効率向上、重大欠陥検出の面で高い評価が得られ、定性評価においても「重大欠陥を効率よく検出できる手法」、「レビュー対象が膨大な場合に有効な手法」という意見を多く得られた。

このことから、自由手法と比較してHDR法は時間制約のある中で重大欠陥を素早く検出するという点で一定の効果が得られる有効な手法であることの確信が得られた。

### 4.2 属人性排除に関する考察

アンケートの結果、「スキル向上に役立つ」という項目で89%の人から有益であるという高い評価を得られた。本手法を使うことによるスキルアップへの期待の高さの現れと言える。

しかし、実践可能性の評価は高いとは言えない結果であった。特に今回の実験で効果の出なかった被験者にその傾向が強い。「やり方の違い」の評価も比較的高いため、各自の従来手法との差異から適用自体に抵抗感があつたことが推察できる。この点は定性評価項目の、「兆候から仮説を導

く部分が難しく、知識や慣れが必要」という意見とも整合するため、「知識移転の容易性」という点で課題が浮き彫りになった。

この課題については、適切な「欠陥知識移転の介助」と兆候-仮説-検査観点のパターン共有等を通じた現場適用時の支援充実が必要である。定性評価で得られた「なぜ欠陥が混入したのか、その原因によって他にどんな欠陥が混入するのかを想起することは、レビュースキル向上に非常に役立つ」という点もその必要性を裏付けている。2.2 章で述べた通り、日頃のレビュー時における、こまめな兆候と欠陥の記録、およびレビュー後の振り返りを継続することで、長期的には「欠陥」や「兆候」等の資産化が実現できる。

以上の内容を総合的に評価すると、手法の導入初期は欠陥知識不足や仮説的思考の不慣れが原因で欠陥検出効率が上がらないが、経験を積み知識を蓄積すれば効率的・効果的なレビュー実施が可能な手法であると評価されたと言える。換言すると、本手法は、従来経験者が少なく体系的に実施されていなかった思考プロセスを用いていること、および継続的实施によるレビューアのスキル向上により生産性・品質の同時向上が可能な手法であると言える。一度目のレビューから結果を出すことは難しいものの、HDR 法という知識のフィードバックサイクルに「欠陥」という知識を乗せて蓄積／共有・移転を行うことで、レビューの検出速度・精度の両面で向上を期待できると考える。

なお、今回は思考メカニズムを形式知化したのが、更なる研究として、本手法がより効果を発揮するためには他にどのようなことが影響するのか、仕様書・ソースコードなどのレビュー対象による違いや、レビューアの専門分野や立場、時間制約の度合いなど様々な要素をパラメータとした実験、また HDR 法を繰り返し実施した場合と通常のレビューを繰り返し実施した場合とでレビュー効果がどのように推移していくのかの比較実験を行い、HDR 法が属人性排除に向けたレビュースキル向上にどれほどの早期実現の効果が期待できるのかを確かめていきたいと考える。

## 5. まとめ

本研究では、トップレビューアの欠陥検出メカニズムに着目して、兆候を基にした「HDR 法 (Hypothesis Driven Review Method: 仮説駆動型レビュー手法)」を提案した。本手法は品質要求・品質特性やレビュー対象物そのものから欠陥の兆候を掴み、その兆候を基に仮説を立てて、狙いを定めて欠陥を検出する、効率的かつ効果的でレビューアのスキル向上にもつながる手法である。実験によって、本手法を用いることで時間制約がある中でも、重大な欠陥を集中的に検出できるようになることの確信を得た。

近い将来、システムの複雑化により網羅的なレビュー実施がますます困難になることを考えると、短時間でも重大な欠陥を検出可能な本手法を活用する場面は多くなると考える。そのためにもいち早く本手法を取り入れ、欠陥知識を蓄え、レビュースキル向上を図り、更なるトップレビューアを育ておくことが必要と考える。

## 参考文献

- [1] 「認知心理学 知のアーキテクチャを探る」有斐閣アルマ 道又邇, 北崎充晃, 大久保街亜, 今井久登, 山川恵子, 黒澤学
- [2] 「間接的メトリクスを用いて欠陥予測を行うレビュー方法の提案-欠陥の位置と種類の特定によりレビューの効率と効果を向上-」SQIP シンポジウム 2011 2010 年度 SQIP 研究会 第 3 分科会