

テスト工程の品質管理の実態に関する意識調査

株式会社NTTデータ 技術革新統括本部 システム技術本部 生産技術部
プロジェクトマネジメント・ソリューションセンター
永田 啓悟

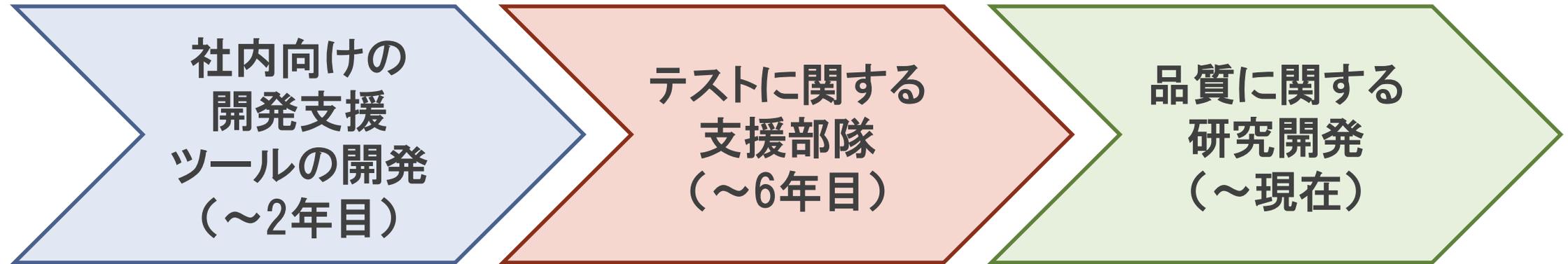
目次

1. 背景
2. 本論
 - 概要
 - スケジュール
 - 活動
 - 仮説立案
 - アンケート作成
 - アンケート依頼
 - 集計・分析
3. まとめ

1. 背景

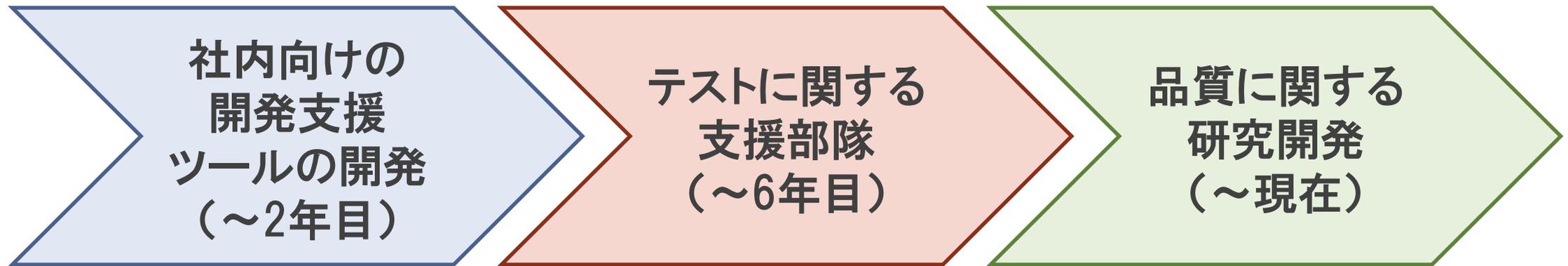
1.1. 自己紹介

- 氏名:永田 啓悟 (ながた けいご)
- テストに関する支援や、品質をテーマとした研究開発に従事。
- 立場上、社内の様々なプロジェクトを見てきた。



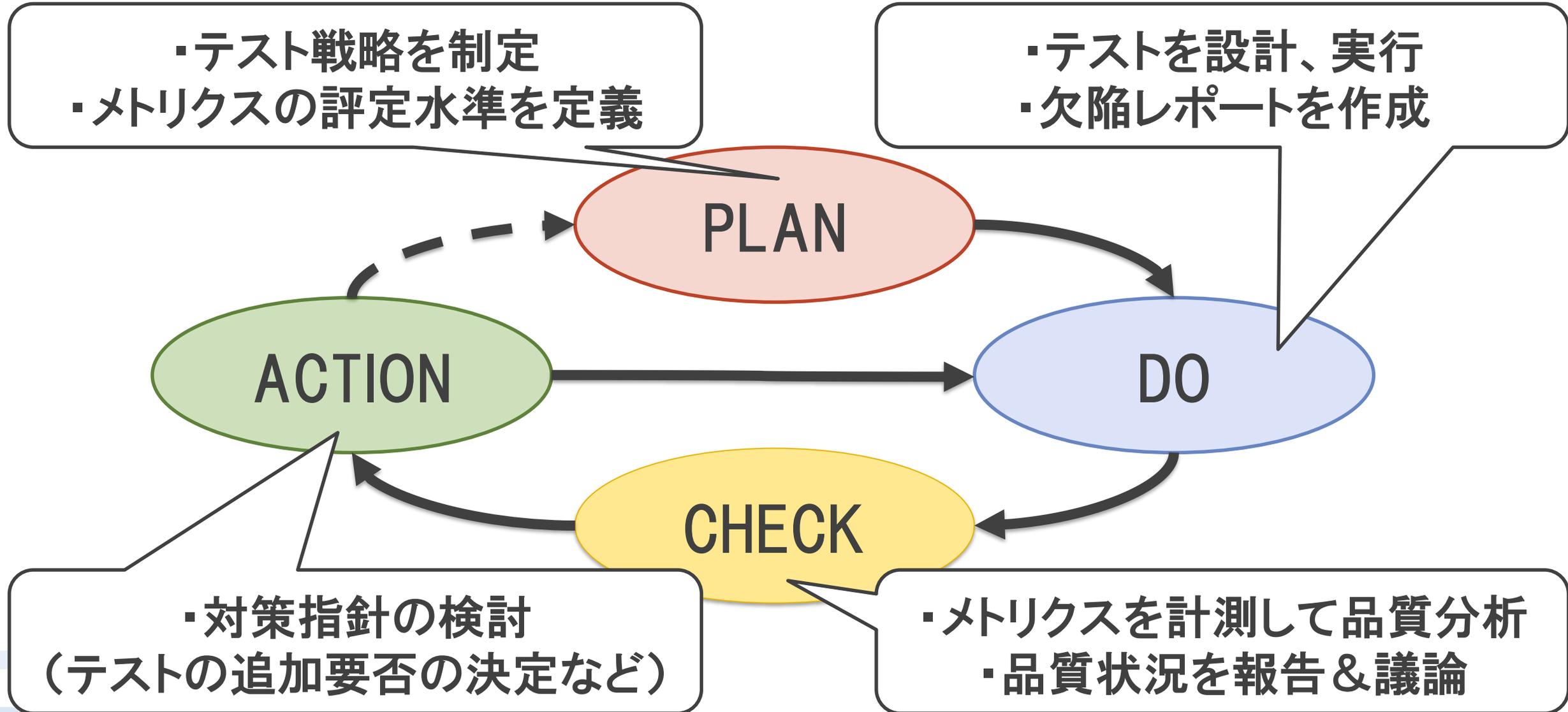
1.1. 自己紹介

- 氏名:永田 啓悟 (ながた けいご)
- テストに関する支援や、品質をテーマとした研究開発に従事。
- 立場上、社内の様々なプロジェクトを見てきた。



品質管理、これで良いのか？

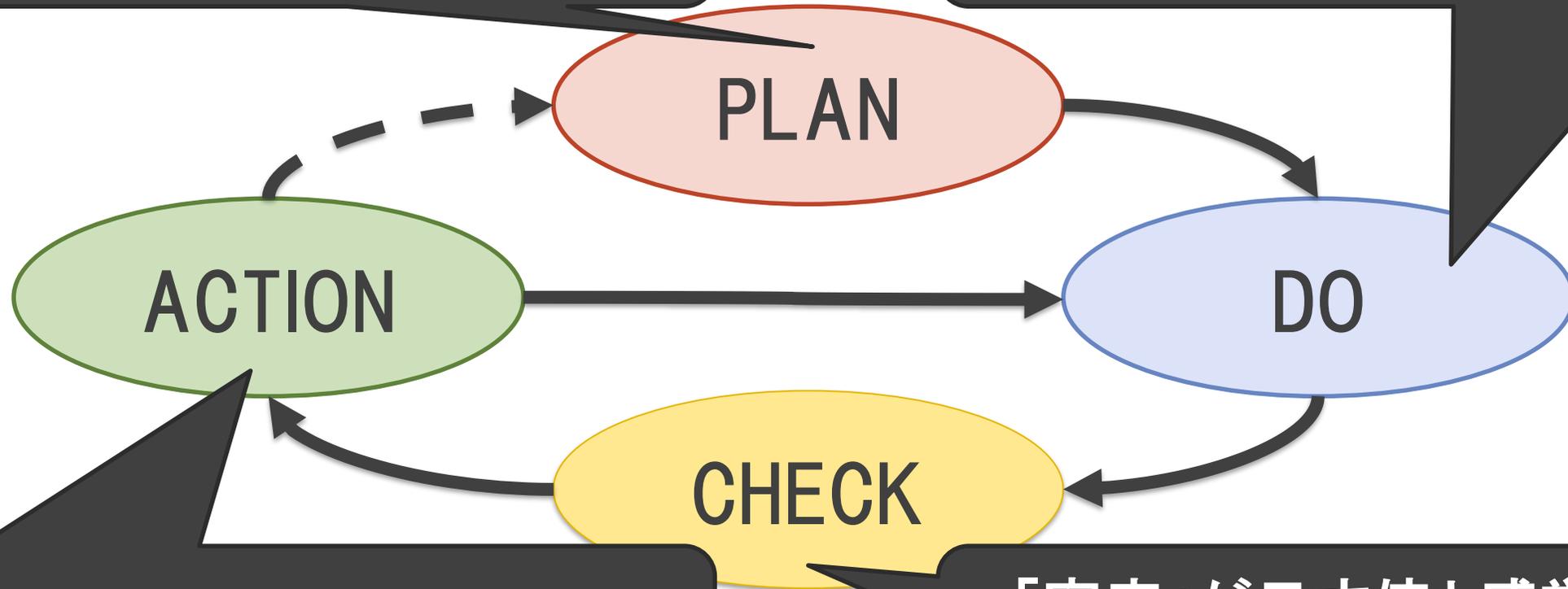
1.2. 当社テスト工程における典型的な品質管理フロー



1.3. その中の、納得できない部分

- ・似たり寄ったりな計画
- ・論拠不明な「密度」の評定水準

- ・バグが出そうにないテスト
- ・機能性確認に偏ったテスト



- ・的を得ていない追加テスト

- ・「密度」が示す値と感覚の乖離
- ・有効活用されない欠陥レポート

1.4. テスト密度と欠陥密度(定義)

$$\text{テスト密度} = \frac{\text{コンポーネント又はシステムに対して実行されたテスト数}}{\text{コンポーネント又はシステムのサイズ}}$$

$$\text{欠陥密度} = \frac{\text{コンポーネント又はシステムの中で検出された欠陥数}}{\text{コンポーネント又はシステムのサイズ}}$$

利用イメージ



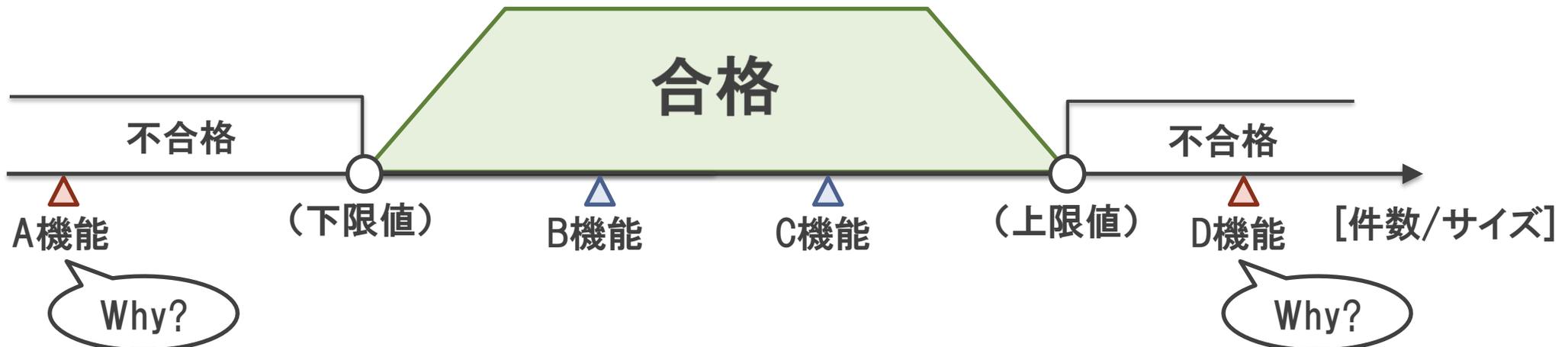
計画時に何かしらの
論理で定義される

1.4. テスト密度と欠陥密度(定義)

$$\text{テスト密度} = \frac{\text{コンポーネント又はシステムに対して実行されたテスト数}}{\text{コンポーネント又はシステムのサイズ}}$$

$$\text{欠陥密度} = \frac{\text{コンポーネント又はシステムの中で検出された欠陥数}}{\text{コンポーネント又はシステムのサイズ}}$$

利用イメージ



1.5. テスト密度と欠陥密度(疑問)

$$\text{テスト密度} = \frac{\text{コンポーネント又はシステムに対して実行されたテスト数}}{\text{コンポーネント又はシステムのサイズ}}$$

$$\text{欠陥密度} = \frac{\text{コンポーネント又はシステムの中で検出された欠陥数}}{\text{コンポーネント又はシステムのサイズ}}$$

上限／下限の設定に用いられる「類似プロジェクト」とは何か？

それぞれの分母／分子の定義はプロジェクト内で統一されているのか？※

これらの値が評価水準を満たすということに、どれほどの意味があるのか？

※社内ルールは定義されています

1.6. 活動の背景

このままで良いのか？

もっと良いやり方があるのでは？

今一度、立ち返ろう

1.6. 活動の背景

このキキで正しいのか？

(その第一歩として)

まずは事実確認＝今日の発表

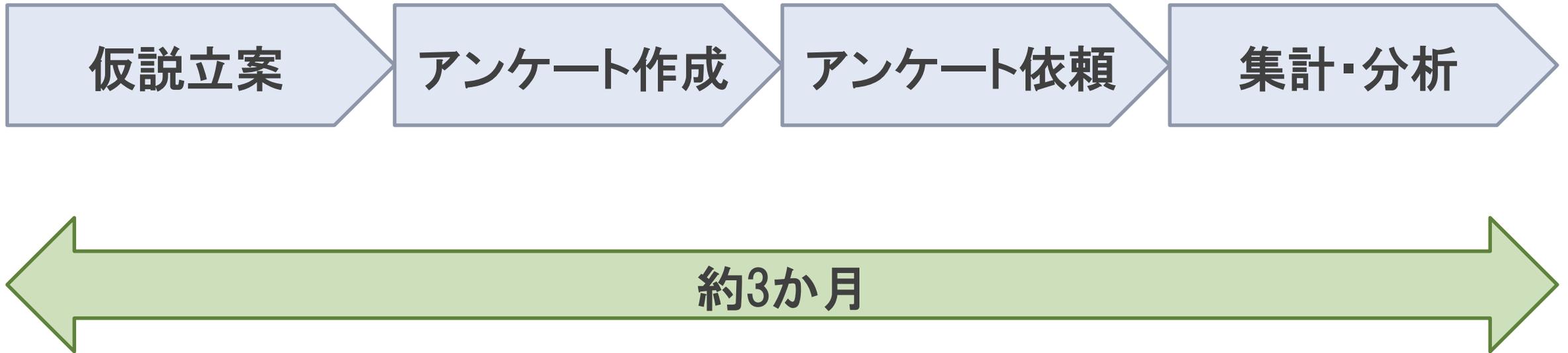
今一度、立ち返ろう

2. 本論

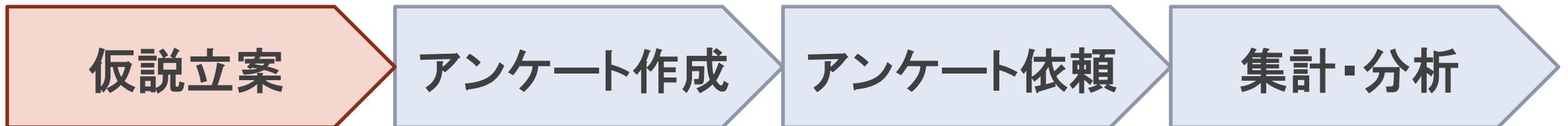
2.1. 今回発表する活動の概要

- テスト工程の品質管理に関するこれらの疑問が、どの程度正しいかを特定するために、アンケートを実施した。
- アンケートの依頼先：
 - 当社
 - 当社のグループ会社
 - 他社(ユーザ系SIer、電機メーカーなど)

2.2. 活動の進め方



2.3. 仮説立案



2.3.1. 仮説立案の方法

品質管理に関する課題や問題を、ブレインストーミングで一覧化。



テスト密度や欠陥密度を用いた説明では、お客様に納得感を与えられていないのではないか。

お客様や製品の特性に応じた、柔軟な品質管理を実施できていないのではないか。

欠陥レポートの執筆やルール浸透に過度な負担がかかっているのではないか。

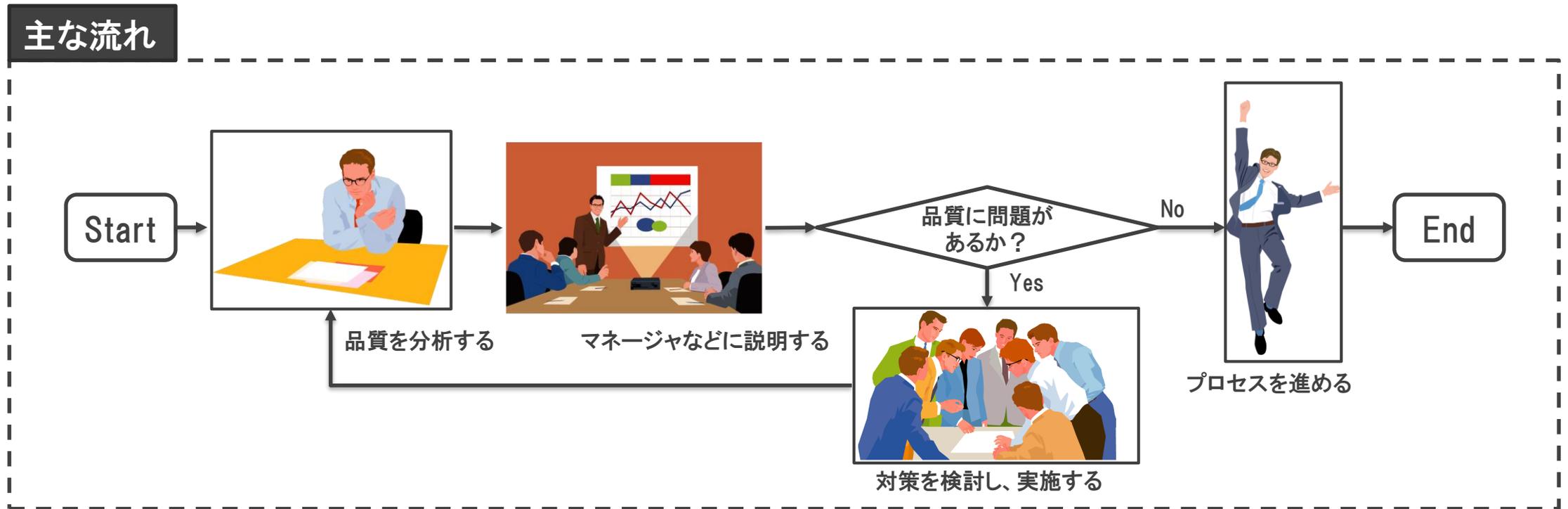
2.3.2. 整理後の仮説一覧(要約)

#	仮説
1	品質評価会議の作業負担が大きい
2	製品やお客様の特性に応じた、柔軟な品質管理を実施できていない
3	欠陥・テスト密度で問題なしと判定した箇所に、後続の工程で欠陥が検出される
4	欠陥・テスト密度を用いた説明では、お客様に品質についてご納得いただけていない
5	欠陥レポート執筆の負担が大きい
6	品質評価会議の効果は負担に見合わない
7	欠陥レポートの記載ルールをプロジェクト内で浸透させることが困難である

#	仮説
8	定量評価の結果に関わらず、定性評価を行って「問題なし」と判定している
9	ローカル環境上の動作テストなど、メトリクスで管理されない活動が行われている
10	品質評価活動がコストの高騰を招いている
11	品質評価活動が開発者や管理者の負担となり、開発作業の停滞や遅延を招いている
12	メトリクスを目標に到達させるためだけに、バグ検出率の低い強化テストを実施している
13	機能や性能などは保証できているが、保守性や満足性などを保証できていない

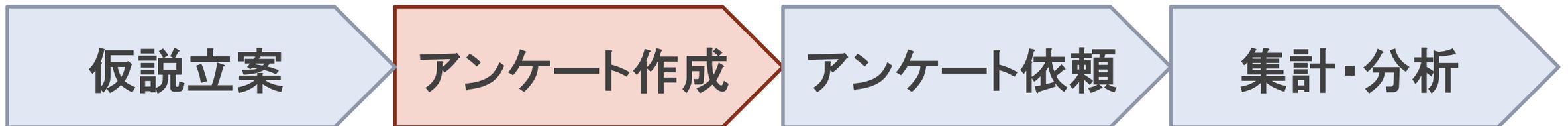
2.3.3. 補足:品質評価会議とは

プロジェクトの特定地点(※)において、品質の分析結果を内部共有し、関係者間で進捗是非の合意を取る会議である。



(※) 「工程終了時」「工程の中間地点」「月次」など

2.4. アンケート作成



2.4.1. アンケートの前提条件

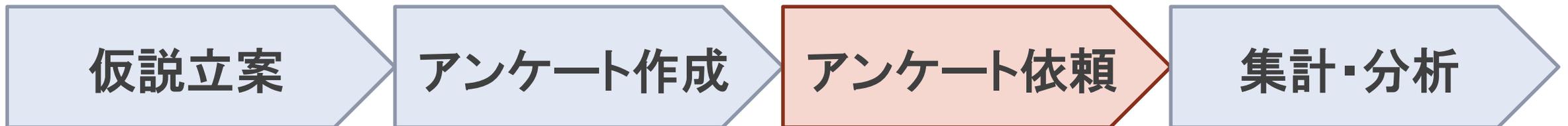
回答に当たっては、**直近で参画していたプロジェクト1件**を思い浮かべて回答してもらうこととした。

- ✓ ただし、「そのプロジェクトでは答えづらい」などの事情がある場合は、任意の1プロジェクトを選定することを許容した。
- ✓ 以後、上記で選ばれたプロジェクトを「**対象プロジェクト**」と呼ぶ。

2.4.2. アンケートの質問内容

#	質問項目	質問内容
1	ご自身について	所属、連絡先など
2	対象プロジェクトについて	対象プロジェクトのミッションクリティカル性、開発期間など
3	品質管理方法について	対象プロジェクトにおけるメトリクスの定義方法や時期、活用場面など
4	品質特性について	対象プロジェクトで保証できた製品品質および利用時品質
5	品質管理活動におけるご経験	品質管理上で想定される問題・課題について、経験・目撃の有無
6	終わりに	意見や感想など

2.5. アンケート依頼



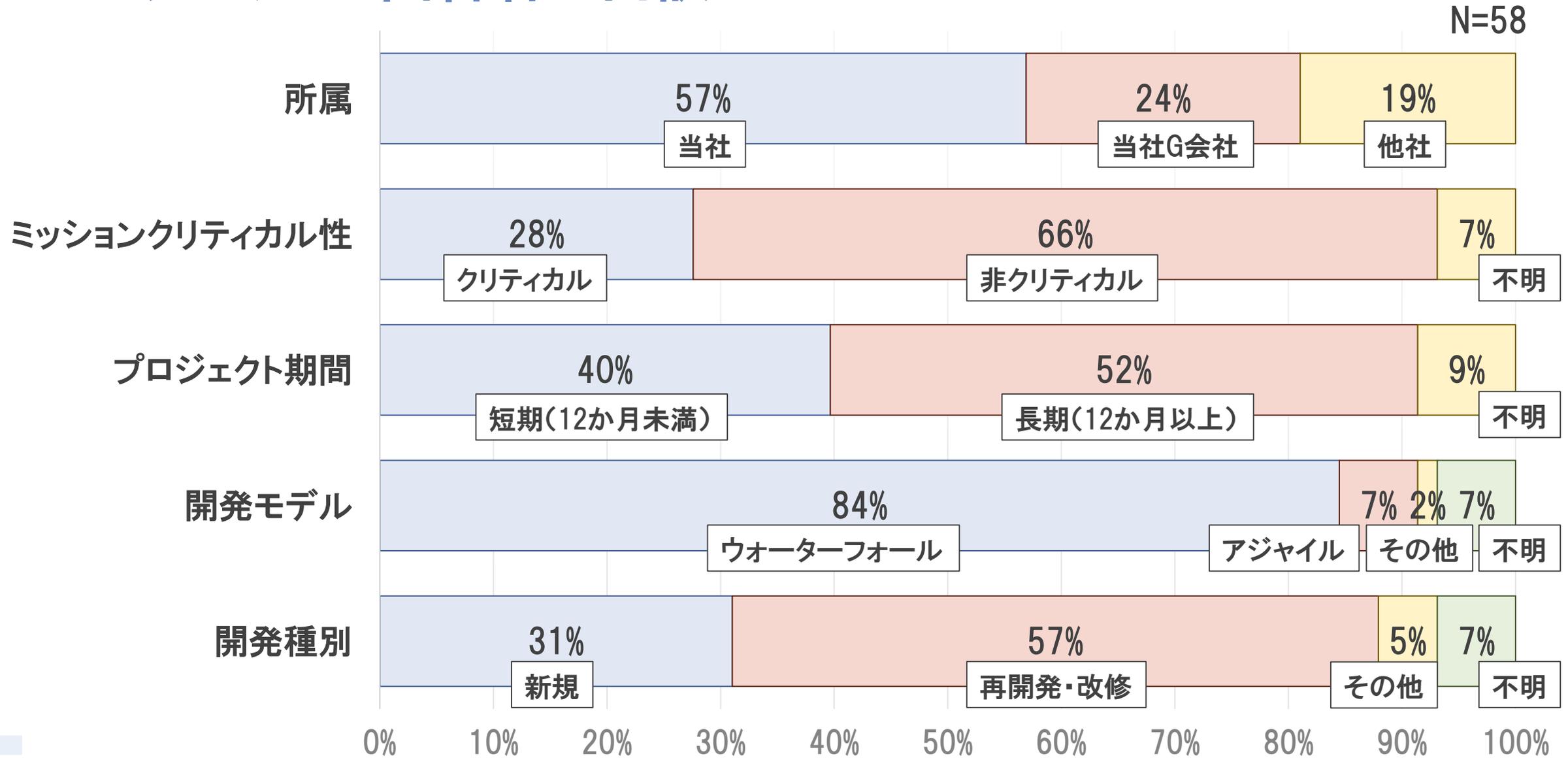
2.5.1. アンケートの配布方法

回答を見込める担当者に、施策の趣旨やアンケート内容を対面で説明し、アンケートのフォーマットを提供した。

その際、周囲に同じく回答していただけそうな方がいれば、そちらへの二次配布も依頼した。

- ✓ 我々から直接配布したのは33名。
- ✓ 二次配布数は不明。
- ✓ 最終的に**58名**から、ご回答を受領した。

2.5.2. アンケート回答者の内訳



2.5.3. 内訳の補足

N=58

所属

IPAの研究会報告書内の記述を参考にしつつ、対象プロジェクトについて以下を確認。いずれかにYesと答えた対象プロジェクトを「クリティカル」、それ以外を「非クリティカル」と定義した。

ミッションクリティカル性

【アンケート】

対象プロジェクトについて、該当するものをご選択ください。(全て選択)

人命に関わる。

停止や誤動作が、社会全体に甚大な経済損失をもたらす。

社会的に、災害時であっても停止が許されない。

プロジェクト期間

JUASの調査報告書内の「開発種別」の分類を借用した。

開発モデル

【アンケート】

対象プロジェクトについて、開発種別をご選択ください。(1つ選択)

新規

再開発・改修

その他

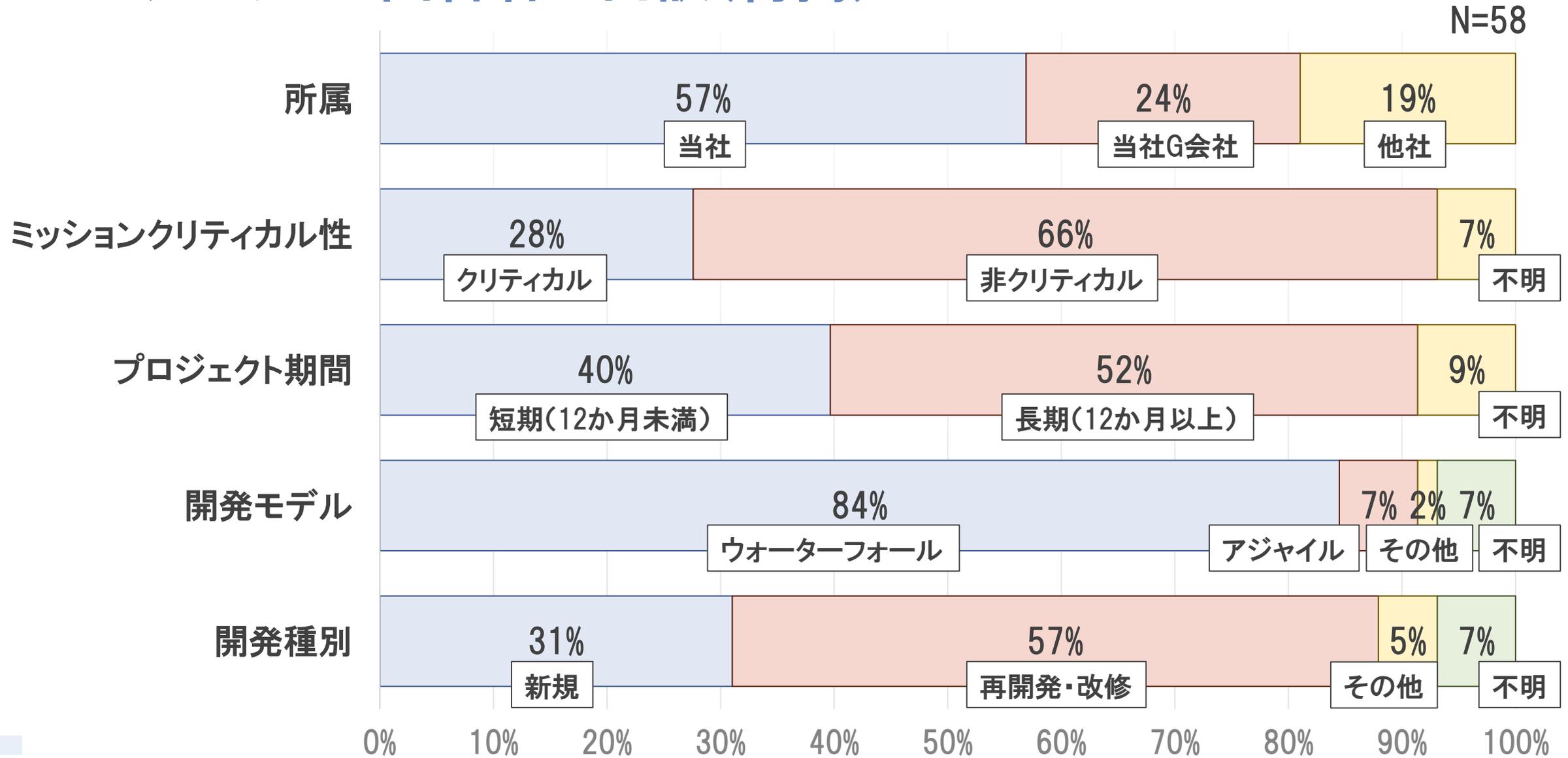
開発種別

0% 10% 20%

出典: <https://www.ipa.go.jp/sec/softwareengineering/reports/20090409.html>

出典: 日本情報システム・ユーザー協会 ソフトウェアメトリクス調査2016

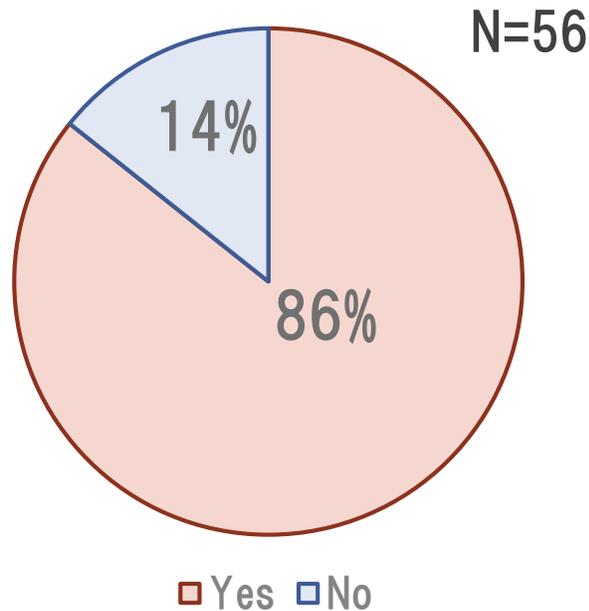
2.5.3. アンケート回答者の内訳(再掲)



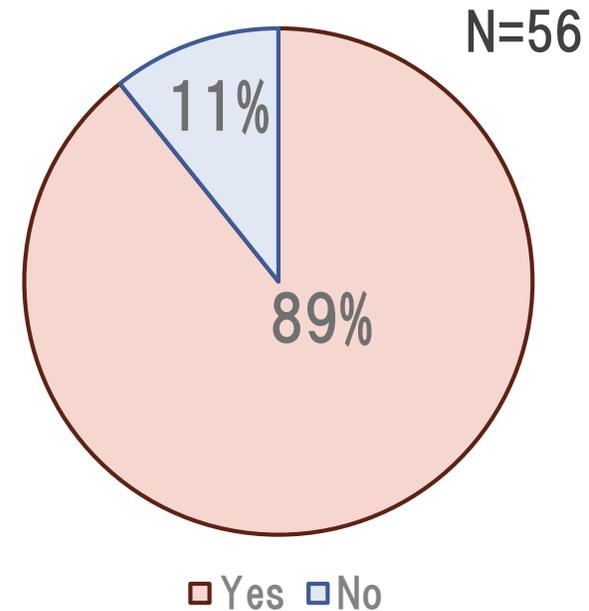
2.5.4. 品質管理プロセス① ～密度の活用状況～

大半の対象プロジェクトにおいて、欠陥・テスト密度の双方が活用されていることを確認した。

テスト密度を使っていましたか



欠陥密度を使っていましたか

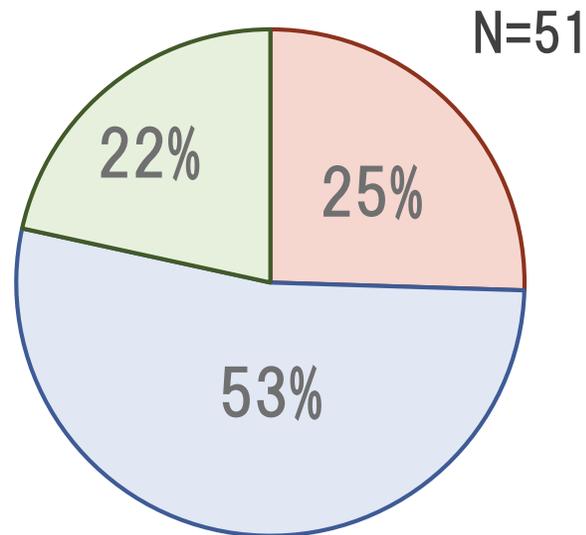


「No」を選択した回答者のコメント：
『テスト数・バグ数などの絶対数』『テスト消化率』など

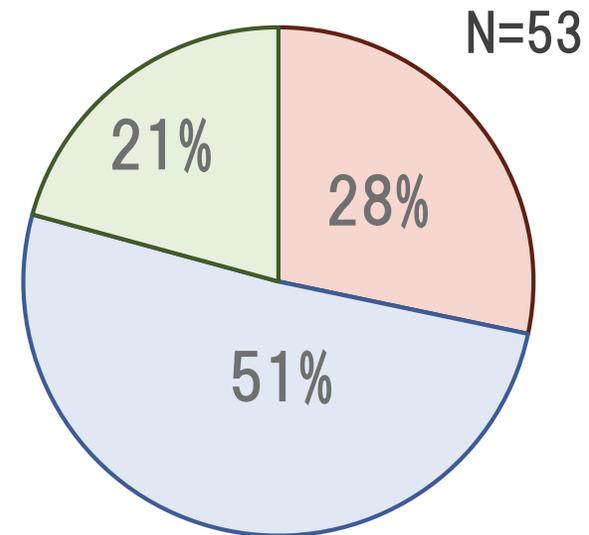
2.5.5. 品質管理プロセス② ～評価水準(上限・下限)の設定方法～

認識どおり、「類似プロジェクトの実績」を拠り所としているプロジェクトが多いことを確認した。

評価水準の設定方法(テスト密度)



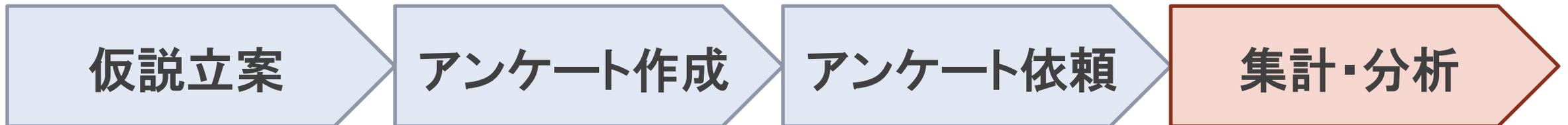
評価水準の設定方法(欠陥密度)



□組織内の指定値 □類似プロジェクトの流用 □その他 □組織内の指定値 □類似プロジェクトの流用 □その他

「その他」を選択した回答者のコメント:
『社外レポートから』『自プロジェクトの過去実績』など

2.6. 集計・分析



2.6.1. 用語定義:同意率

各仮説に対し、肯定的なアンケート回答の選択率を「**同意率**」と呼ぶ。

仮説

品質評価会議の効果は負担に見合わない



質問

品質評価会議にかけた負担に見合う効果を感じますか



この場合は、「No」の選択率を、この仮説に対する「同意率」と定める。

2.6.2. 同意率のランキング(50%以上)

順位	仮説	同意率
1	品質評価会議の作業負担が大きい	87%
2	製品やお客様の特性に応じた、柔軟な品質管理を実施できていない	71%
3	欠陥・テスト密度で問題なしと判定した箇所に、後続の工程で欠陥が検出される	64%
3	欠陥・テスト密度を用いた説明では、お客様に品質についてご納得いただけていない	64%
5	欠陥レポート執筆の負担が大きい	55%
6	品質評価会議の効果は負担に見合わない	52%
7	欠陥レポートの記載ルールをプロジェクト内で浸透させることが困難である	50%

2.6.3. 同意率のランキング(50%未満)

順位	仮説	同意率
8	定量評価の結果に関わらず、定性評価を行って「問題なし」と判定している	36%
9	ローカル環境上の動作テストなど、メトリクスで管理されない活動が行われている	33%
10	品質評価活動がコストの高騰を招いている	28%
11	品質評価活動が開発者や管理者の負担となり、開発作業の停滞や遅延を招いている	26%
12	メトリクスを目標に到達させるためだけに、バグ検出率の低い強化テストを実施している	17%
—	機能や性能などは保証できているが、保守性や満足性などを保証できていない	次頁参照

2.6.4. 同意率で表現できない仮説の結果について①

仮説

機能や性能などは保証できているが、保守性や満足性などを保証できていない



質問

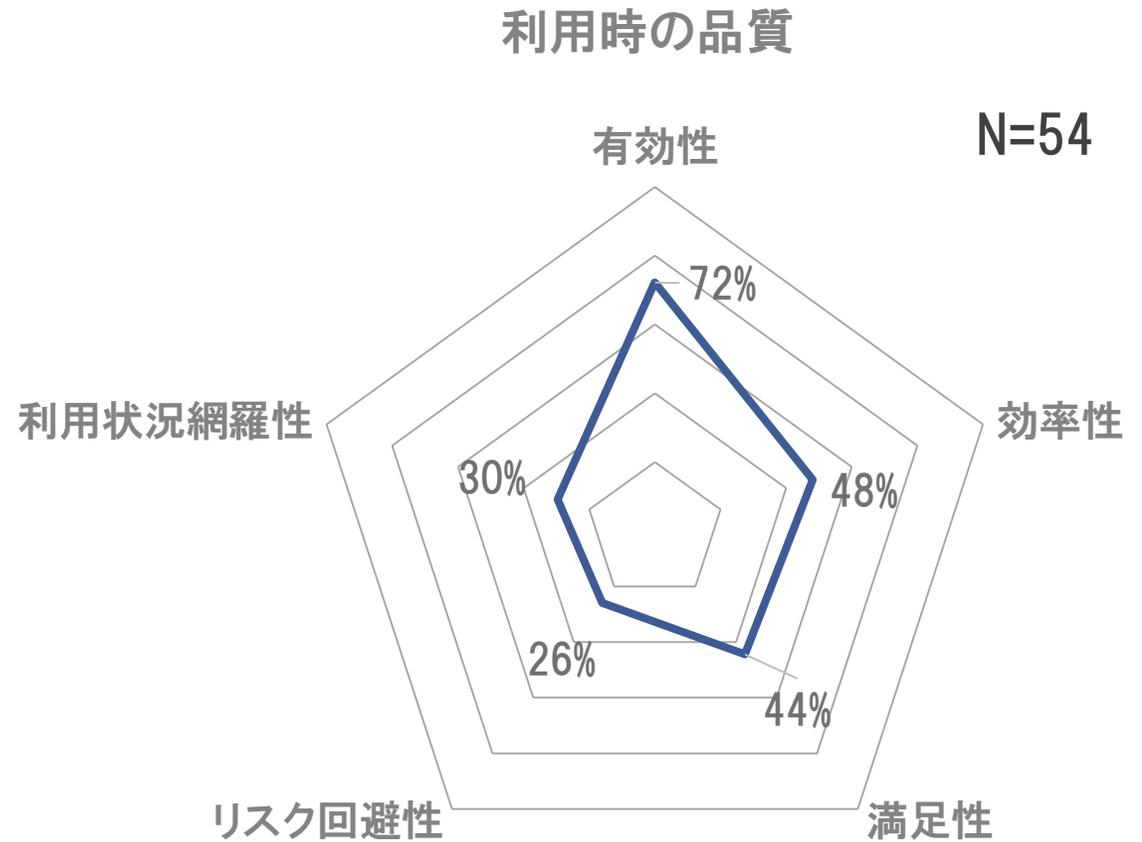
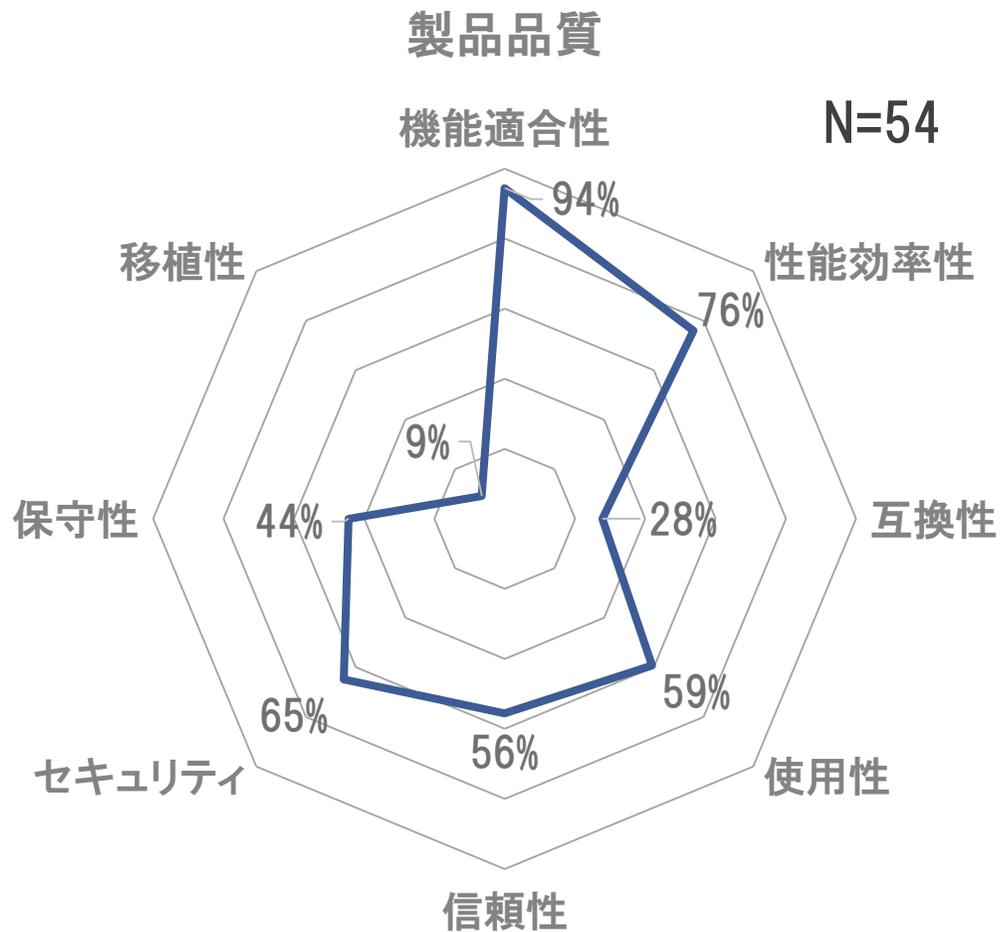
対象プロジェクトで品質が確保できていたと考えている品質特性をお聞かせください。

- 機能適合性
- 性能効率性
- 互換性
- 使用性
- 信頼性
- セキュリティ
- 保守性
- 移植性
- 有効性
- 効率性
- 満足性
- リスク回避性
- 利用状況網羅性

品質特性ごとのYes回答率を集計し、偏りを確認する

出典: <http://kikakurui.com/x25/X25010-2013-01.html>

2.6.5. 同意率で表現できない仮説の結果について②



CAUTION

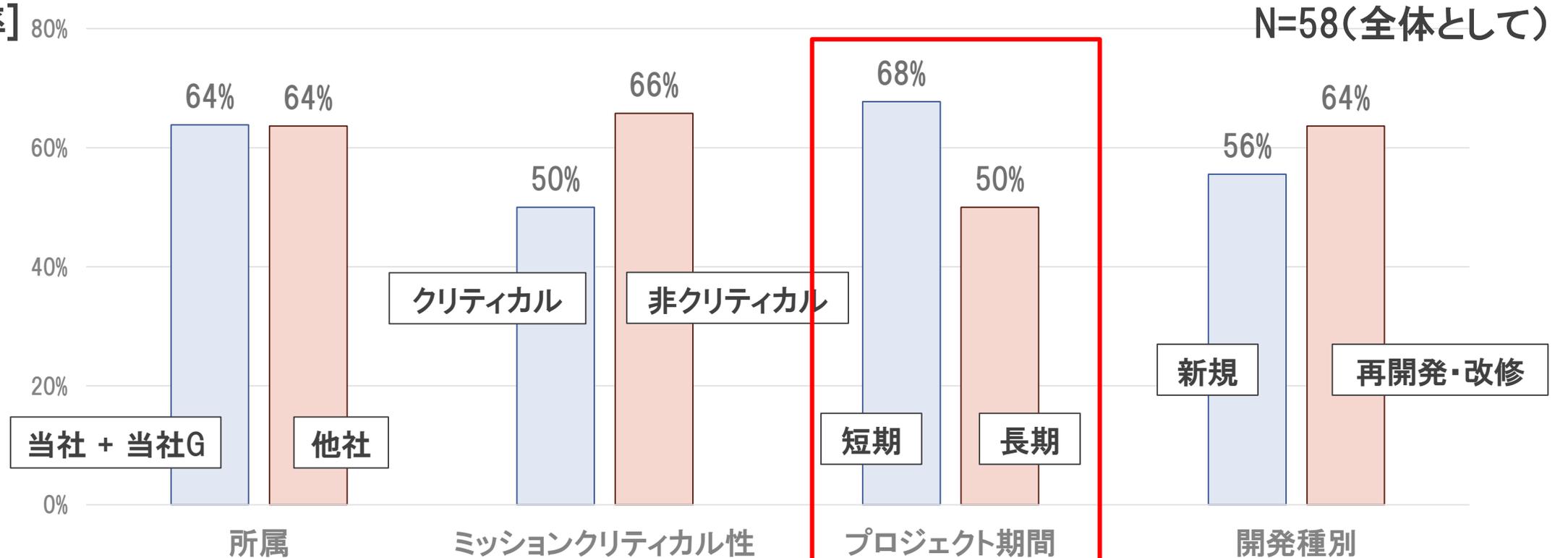
「値が低い」≠「その特性について当社システムに問題がある」

2.6.6. 仮説のドリルダウン①

仮説

欠陥・テスト密度を用いた説明では、お客様に品質についてご納得いただけていない

[同意率] 80%

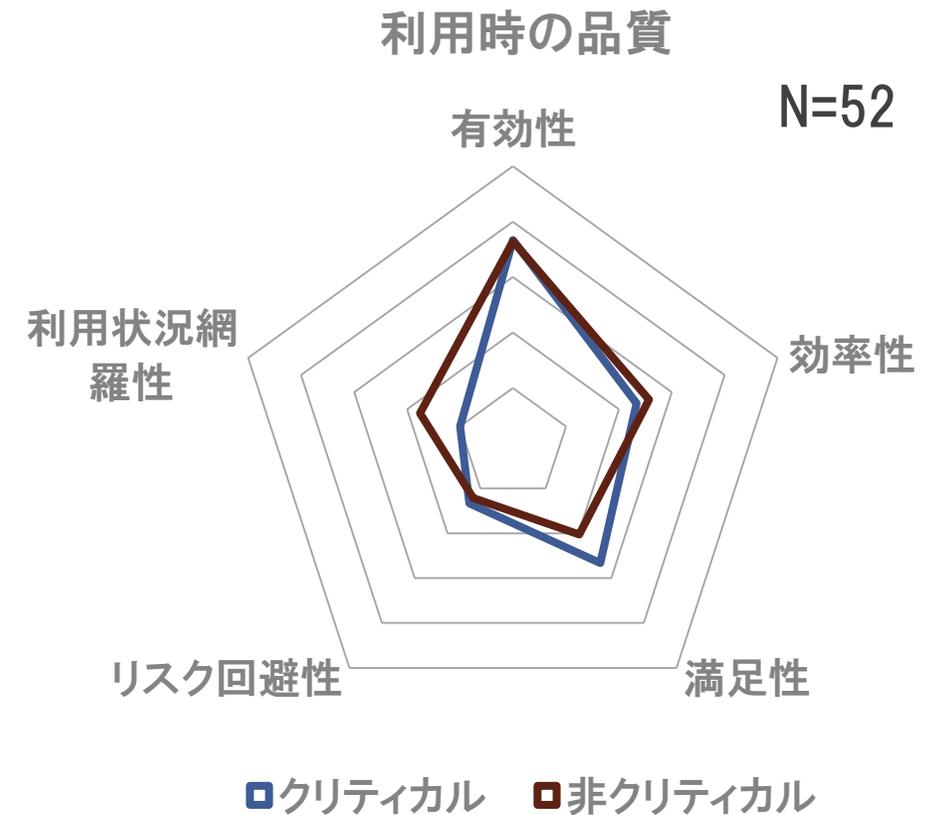


「非同意」を選択した回答者のコメント:
『密度による説明はお客様指定』など

2.6.7. 仮説のドリルダウン②

仮説

機能や性能などは保証できているが、保守性や満足性などを保証できていない



2.6.8. 仮説のドリルダウン③

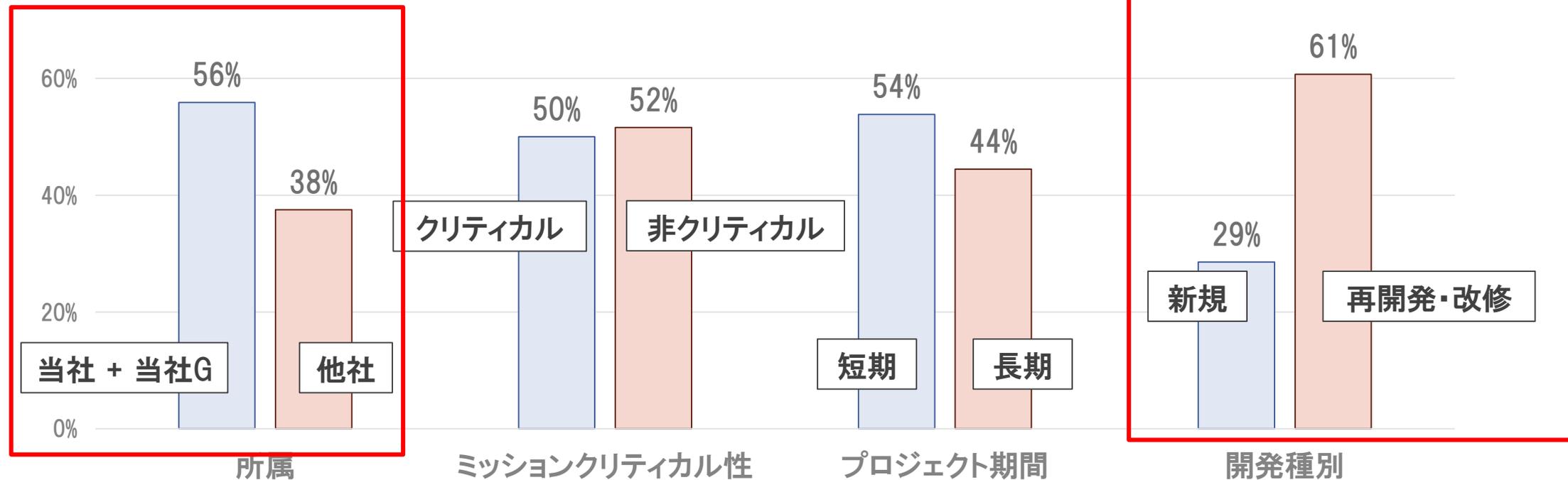
仮説

品質評価会議の効果は負担に見合わない

[同意率]

80%

N=47(全体として)



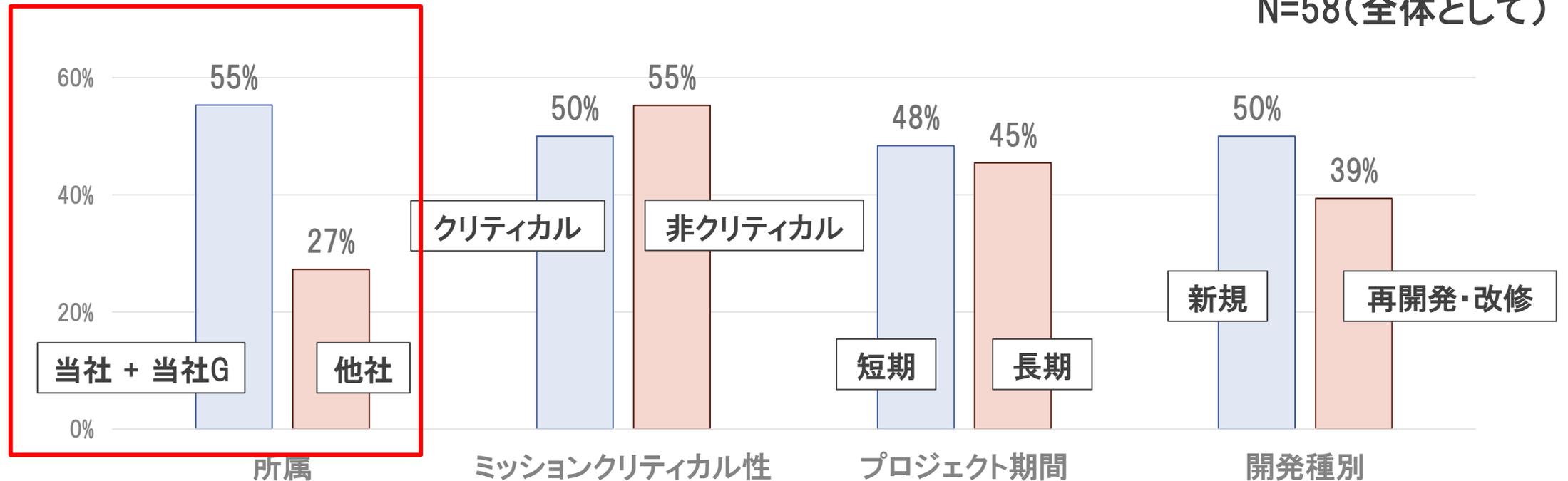
「非同意」を選択した回答者のコメント：
『品質強化策を打つための最後の防波堤にはなっている』など

2.6.9. 仮説のドリルダウン④

仮説

欠陥レポートの記載ルールをプロジェクト内で浸透させることが困難である

[同意率] 80%



3. まとめ

3. まとめ

テスト工程の品質管理に関するアンケートを実施、58件の回答を受領。

想定内

- ✓ 現場の品質管理の中心は「密度」。
- ✓ お客様特性・プロジェクト特性を品質管理に反映できていない。
- ✓ 品質評価会議の負担に対する効果の薄さ。
- ✓ 当社欠陥レポートの難解性。



- ✓ あるべき品質管理の検討。
- ✓ 当社固有課題については、他社様の良いところの分析、吸収。

想定外

- ✓ 否定に偏った仮説も多い。
- ✓ 特に追加テストに関する、「手段と目的の逆転」という仮説は強く否定された。



- ✓ さらなる要因調査。

共通アクション

さらなるアンケートの収集(特にアジャイル、他社)



NTT DATA

Trusted Global Innovator