

製品担当者によるユーザ現場観察の有効性評価

小幡 明彦¹ 高本 康明² 木島裕二¹ 國藤 進³ 杉山公造³
(株)富士通研究所¹ (株)富士通² 北陸科学技術先端大学院大学 知識科学研究科³

あらまし

顧客中心の製品デザイン手法として、文脈における質問法というユーザ観察手法が関心を集めている。製品開発の担当者によって、ユーザが業務を遂行している現場に訪問し、業務の最中に割り込んで質問するなどの方法でインタビューを実施し、観察結果をもとにワークモデルの記述やKJ法を実施することによって製品アイデアを抽出する。しかし、製品担当者にインタビューの方法や、ワークモデルの分析方法を訓練する必要があり、どれだけ効果的にユーザ観察が実施できるのか明らかにされていない。本稿では、必要最小限の訓練で実施可能な文脈における質問法簡略版を提案し、従来のフォーカスグループとの比較により、その有効性を評価する。

Evaluate the effectiveness of a field observation conducted by a product team

Akihiko Obata¹ Yasuaki Takamoto² Yuji Kijima¹
Susumu Kunifuji³ Kozo Sugiyama³
Fujitsu Laboratories Ltd. ¹ Fujitsu Limited ²
Japan Advanced Institute of Science and Technology³

ABSTRACT

Contextual Inquiry method has been currently receiving increased attention as a method of customer focused design. Contextual Inquiry gets data from users by having product engineers interview users in their workplace as they do work, and then produces design ideas by analyzing work models and affinity diagrams. However, it requires engineers to be trained in interview techniques. Moreover, the effect of Contextual Inquiry is still unknown. In this paper, we proposed a discount method of Contextual Inquiry that requires minimal training, and evaluates the effectiveness of the method comparing with Focus Groups method.

1. はじめに

1990年代より、グループウェアシステムの設計方法として、エスノグラフィー (Ethnography) をベースにしたユーザ行動の調査方法が注目を集めており、航空機の管制塔の設計などへの適用例が CSCW などの学会で報告されている[1]。しかし、これらの方法では、エスノグラフィーの専門家の成果を、開発者が理解しないことや、短時間に結果を出すのが困難であるなどの問題点が指摘されている[2]。

これに対して、Holtzbratt 等によって提案された文脈における質問法 (Contextual Inquiry) [3,4] は、製品の開発者、販売推進などの職能の横断チームでユーザの現場に訪問し、10数名程度のユーザを、それぞれ数時間だけインタビューを実施し、観察結果を横断チームで分析するなど、より実践的な手法である。しかし、製品の担当者自身が現場観察を行うことで、フィールドワークの担当者と製品開発者との意見の衝突をさけることができるが、製品担当者が、インタビューの方法やワークモデルの作成など観察データの分析方法を習得する必要がある。また、どの程度の訓練をすれば、短時間の観察で有効なデータが得られるのか、明らかにされていないことも事実である。

Curtis 等は、文脈における質問法の実施について詳細に報告している[5]。50%以上の工数をさけることを前提にチームメンバを募集し、3ヶ月にわたるユーザの現場訪問によって、1500の観察データを収集した。非常に多くのアイデアの抽出に成功したが、実際に開発部門がそれをどの程度役立てることができたかを追跡することができなかつたと報告している。小規模なプロジェクトや、ライフサイクルの短い製品のプロジェクトでは、このような不透明な成果に対して、これだけの人材や工数を確保することは非常に困難である。

本稿では、現場観察の経験がないチームが、A4一枚の簡単な手引きと、30分以内の簡単な説明だけで実施可能な文脈における質問簡略版 (Discount Contextual Inquiry) を提案し[6]、現

在広く実施されているフォーカスグループ[7]による観察結果と比較することで、その有効性を検証する。まず、2章では、文脈における質問簡略版の手順を説明し、3章では、その有効性の評価方法について説明する。4章では、評価結果について説明し、5章で、今後の課題を述べる。

2. 文脈における質問簡略版の実施手順

文脈における質問を製品開発プロセスへ導入するまでの、躊躇、抵抗感を減少させるには、訓練にかかるコストを最小限におさえることが重要である。簡略版では、インタビューの訓練は、A4一枚の簡単な資料で30分以内で説明するだけであり、観察データの分析は、ワークモデルによる分析を省略し、普段の会議とかかわらない方法で作業を行う。ここでは、文脈における質問簡略版[6]の観察方法、および、分析方法の概要を述べる。

(1) 観察方法

製品担当者が訓練なしにインタビューを行えば、日頃気にしている疑問をユーザにぶつけ、誘導的な質問に終始するおそれがある。少なくともインタビューの方法に関する最小限の知識が必要である。そこで、30分程度で説明できるインタビューの手引きを作成した[6]。インタビューは、基本的には、ユーザのアクションをきっかけにして、アクションに関する質問をする。例えば、ユーザが他者とコミュニケーションをした直後や、書類を受け渡しした直後、システムを操作した直後等に、何がきっかけで、何の目的で実施したのか、また、授受した情報の意味、構造などを5W1Hの形式で質問をする。さらに、ユーザの説明に対して、それを、詳細化、あるいは、一般化していく質問をぶつけていく。アクションをきっかけにした質問の生成は、ある意味で機械的な作業であり、訓練や経験がなくても容易に実施することができる。

ユーザに許された短い観察時間 (2時間程度) では、製品と関係する業務がほとんど行われな可能性もある。そこで、業務で用いる書類や機器などのアーティファクトを見せてもらい、そのアーティファクトに対してなされたアクションの結

果（書類に書かれた内容や、ログデータなど）をきっかけにして、どのような状況でそのアクションを実施したかを質問するインタビューをおりまぜて実施する。

（2）観察データの分析

現場観察を実施したチームで、現場観察直後に、観察した事実の報告会を行う。普段の会議における議事録作成と同じ要領で、報告された事実、質疑応答を簡条書きに作成する。また、複数の地点で得た観察結果をもとにKJ法[7]を実施し、観察結果のまとめを作成する。

3. 評価方法

文脈における質問簡略版の有効性を、専門家によるフォーカスグループとの比較により検証する。両手法のパフォーマンスの比較は統制実験では困難であるため、実際の実践活動を比較することにした。市場投入から数年経過している実際の携帯情報機器に対して、次期バージョンに対する製品アイデア、販売促進のアイデアを得ることを目的として、2つの手法でユーザ観察を実施し、そのアウトプットとなる観察結果に対して、主観評価、および、アイデア抽出作業のパフォーマンスを比較する。観察の対象となるユーザの業務、人数、観察時間は、それぞれの条件で共通となるようにした。

3. 1. 文脈における質問簡略版の実施手順

観察チームは、携帯情報機器のソフトウェア開発部門、販売推進部門、及び、デザイン部門、研究部門から現場観察の経験がない6人のメンバによって構成した。観察対象のユーザとして、ターゲットとなっている携帯情報機器の導入を予定している3地点の営業所の営業職10名が参加した。各現場では、インタビューの手順を30分程度説明し、1名のユーザに対して、1名の観察者を割り当て、オフィスでの業務観察と、回想型インタビューを含めて2時間行った。現場でのユーザの観察後、観察結果の報告会をユーザの現場の会議室で、2時間実施した。ユーザの観察事実、ユー

ザの生のコメント、観察者自身のコメント（アイデア、課題など）を区別して、50字程度の簡条書きにして記録した。3地点の現場観察終了後、各地点で記録した139件のユーザの観察事実、及び、ユーザのコメントをもとにKJ法を実施し、139件の生の観察データから、200字程度の分析結果を4件作成した。

3. 2. フォーカスグループの実施手順

フォーカスグループのチームは、司会、書記、観察者の3名の社外の専門家によって構成した。フォーカスグループは、ターゲットとなっている携帯情報機器の最新バージョンを実際に利用している4名のユーザと、他社製品を利用している6名のユーザ、計10名を対象に実施した。10名のユーザを、5名の2組のグループにわけ、各グループに対して、それぞれ、2時間、文脈における質問の条件と共通の業務にフォーカスを設定し、その時のアプリケーションの利用状況について議論を実施した。

フォーカスグループのチームは、ユーザのコメント、及び、分析結果を、文脈における質問のチームと同様に、それぞれ50字、200字程度で報告した。その結果、82件のユーザコメント、6件の分析結果が報告された。

3. 3. 観察結果の比較手順

文脈における質問、及び、フォーカスグループで得られた観察データ（ユーザの生のコメント、および、観察事実）と、各チームが報告した分析結果に対して、製品開発担当の7名による主観評価に加え、それをもとにして実際にアイデアが抽出できるかどうかのパフォーマンスを比較評価した。

観察データに関しては、文脈における質問、フォーカスグループで、それぞれ、139件、82件と大量にあるため、2名のユーザビリティの専門家が重要度づけを独立に実施し、それぞれから重要度の平均値の高い上位30件の観察データを抽出した。分析結果については、文脈における質

間で報告された4件と、フォーカスグループで報告された6件のうち重要度の高い上位4件を評価対象とした。

文脈における質問、及び、フォーカスグループの観察データ、および、分析結果を、それぞれを一つのリストにランダムな順にならべ、提示されたデータがいずれの実験結果の報告であるかわからないようにした。7名の製品担当者は、それぞれ異なる順番にソートされた観察データ、及び、分析結果に対して、アンケートへの回答、及び、アイデア抽出を行った。

3. 4. 現場観察に期待する効果

文脈における質問簡略版などによる現場観察は、ユーザの許可を得る必要があるが、ユーザの現場から離れて実施されるフォーカスグループと比較して次のような効果が期待できる。

(1) 意外性

文脈における質問簡略版は、製品の印象よりもユーザの業務の観察に重点をおいている。製品そのものの印象に重点をおいたフォーカスグループよりも、製品担当者にとって意外で新鮮なデータを得やすい。

(2) 見過ごされてきた事実の明確化

文脈における質問は、ユーザが無意識に行っている活動で見過ごされてきた事実を、再認識させる効果がある。フォーカスグループでは、普段、無意識に行っている活動については、ユーザの印象に残っていないため、報告されない傾向にある。

(3) 普遍性

フォーカスグループは、現場をはなれた場での発言であるため、その場の雰囲気にもとづく思いつきの発言が多く、文脈における質問よりも普遍性が低いと評価される傾向が高い。

(4) 製品開発における有効性

文脈における質問の方が、製品設計のアイデアにつながるアイデアを得やすい。

(5) 販売推進における有効性

文脈における質問の方が、広告のキャッチフレーズなど販売促進につなげるアイデアを得やすい。

4. 評価結果

表1、表2に観察データ、および、分析結果に対する主観評価、および、表3にアイデア抽出率に対する分散分析の結果を示す。表中、CIは、文脈における質問を、FGはフォーカスグループを示し、それぞれ、アンケートの5段階評価の平均値、分散分析のF値、P値を示す。アンケートは、1が否定的、5が肯定的であることを示す。分散分析は、30件の観察データに対する各被験者の評点の平均値を算出し、1要因被験者内計画のモデルによって実施した。

4. 1. 主観評価

(1) 意外性

「自分にとって新鮮、意外な情報か?」については、観察データについては、期待した通り、文脈における質問の観察データの方がフォーカスグループのそれより意外性の高いデータを有意に多く含んでいることが確認できた。しかしながら、分析結果の報告についての主観評価は、文脈における質問とフォーカスグループの有意差は確認されなかった。文脈における質問の条件では、観察データより分析結果の方が低い評価となっており、観察データを分析することで、観察データの意外性をまるめてしまった可能性があることを示唆している。

(2) 見過ごされた事実の明確化

「設計、販促活動で十分意識していた情報か?」については、観察データに関しては、文脈における質問の方が、フォーカスグループよりも、見過ごされてきた事実を有意に多く含んでいることが確認できた(危険率5%)。しかし、分析結果については、意外性と同様に、有意差を確認することはできなかった。文脈における質問、フォーカスグループの双方ともに、分析結果の方が観察データの評価よりも低い評価になっている。分析結果だけを開発担当者に報告するのでは、せっかくの成果をそこなう可能性があることを示唆している。

(3) 普遍性

「普遍的な情報か？」については、予想に反して、文脈における質問と、フォーカスグループとの差異は、観察データ、分析結果ともに確認されなかった。観察データに対して、評価者が2以下の評価を与える率は、文脈における質問では、18%、フォーカスグループでは、13%であり、むしろ文脈における質問の方が、特殊として評価される観察データが多い。分析結果の評価については、文脈における質問、フォーカスグループともに、観察データよりも普遍性の評価が高い。これは、分析作業の際に、分析者にとって特殊であると評価された観察データを振り落とししてしまったためと考えることができるであろう。

(4) 製品開発における有効性

製品開発における有効性についての主観評価では、観察データ、分析結果ともに文脈における質問の方が高く、観察データに対しては1%の危険率で、分析結果に対しては5%の危険率で有意差が確認された。

(5) 販売推進活動における有効性

販売推進に役に立つかどうかについては、製品開発における有効性とは異なり、文脈における質問とフォーカスグループとの有意差は確認されなかった。販売推進活動に対しては文脈における質問とフォーカスグループの有効性は同等であると考えることもできるが、7名の評価者に販売推進の担当が加わっていないことが原因の可能性もある。販売推進活動での有効性についての結論は保留としたい。

4. 1. パフォーマンス評価

観察データに対するアイデア抽出率は、文脈における質問の方が高く、5%の危険率で有意差が確認された。製品開発における有効性を問う主観評価と、整合する結果となった。しかしながら、分析結果については、文脈における質問とフォーカスグループとの間に有意差は確認されなかった。分析結果から抽出されたアイデア数は、文脈における質問、フォーカスグループともに、観察デー

タから抽出したアイデア数よりも大幅に少ない。なぜ、分析結果に対しては、アイデアを記述できなかったかを評価者にインタビューすると、具体的な観察データに対しては、具体的なアイデアを記述できるが、抽象的な分析結果に対して、抽象的なアイデアを記述しなければならないと思われ、それが非常に困難であった等のコメントを受けた。従って、このデータから、必ずしも分析結果の質が低いとは結論できないが、具体的な観察結果そのものを観察担当者間で共有することが効果的であることを示唆している。

6. まとめ

本稿では、開発担当、販売推進担当など現場観察の経験のないチームがA4一枚の簡単な手引きと、30分以内の簡単な説明だけで、訓練なしにインタビューを実施し、ワークモデルによる分析作業を省くことにより、特別な訓練が不要な文脈における質問簡略版を提案した。製品担当者による文脈における質問簡略版と、専門家の実施したフォーカスグループによって得た観察結果を比較する事例研究の結果、文脈における質問簡略版の方が有効性の高い観察データを得ていることを確認することができた。文脈における質問簡略版の方が、意外性が高く、見過ごされてきた事実を明確化することが可能であり、製品設計に役立つ情報が得られると評価された。また、観察結果に対して、より多くのアイデアを抽出することも確認することができた。インタビューの方法についての専門的な知識をもたない開発者による現場観察が、専門家によるフォーカスグループよりも有効な結果が得られた事例を示すことで、現場観察実施に対する開発者の不安感を抑制することができ、ユーザビリティ担当者のリーダーシップの発揮が容易になるであろう。

本研究のもう一つの成果は、観察データの分析結果をまとめることで、観察データの意外性をまらめてしまう可能性や、アイデア抽出が困難になることを示唆するデータが得られたことである。観察データは膨大であるため、なんらかの抽象化

する手順は必要である。観察データの有効性をそ
こなわずに分析結果を導く方法を検討していく
ことが、今後の課題としてあげられる。

参考文献

- [1] Bentley, R., Hughes, J., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., and Sommerville, I.: Ethnographically-informed systems design for air traffic control, Proc.CSCW'92, 123-129, (1992)
- [2] Hughes, J., King, V., Rodden, T., and Andersen, H., Moving out from the control room: Ethnography in system design, Proc.CSCW'94, 429-439 (1994)
- [3] Beyer, H., Holtzblatt, K.: Contextual design: Defining Customer-Centered Systems; Morgan Kaufmann Publishers, Inc. (1998)
- [4] Holtzblatt, K.: Customer centered design as discipline; Human-Computer Interaction - INTERACT'99 (Sasse, A., Jhonson, C.), IOS Press, 3-17, (1999)
- [5] Curtis, P., Heiserman, T., Jobusch, D., Notess, M., Webb, J.: Customer-focused design data in a large, multi-site organization; Proc. CHI'99., ACM, 608-615, (1999)
- [6] 小幡, 高本, 木島, 國藤, 杉山: 簡略化した手順による文脈における質問法の有効性: 事例研究、ヒューマンインタフェース学会論文誌 (掲載予定)
- [7] 川喜田: 発想法 中公新書 (1967)

表1 観察結果に対する主観評価

	C I 平均値	F G 平均値	F 値	P 値
自分にとって新鮮、意外な情報か	3. 2	2. 8	12. 3	0. 01*
設計、販促活動で十分意識していた情報か	3. 2	2. 8	10. 5	0. 02*
普遍的な情報か	3. 4	3. 5	0. 98	0. 36
製品設計に役立つか	3. 4	3. 1	21. 8	0. 01**
販売推進に役立つか	3. 1	3. 1	0. 09	0. 78

アンケートの回答は、5が肯定的、1が否定的。*は $p < 0.05$ 、**は $p < 0.01$ を示す。

表2 分析結果に対する主観評価

	C I 平均値	F G 平均値	F 値	P 値
自分にとって新鮮、意外な情報か	2. 8	3. 0	1. 83	0. 22
設計、販促活動で十分意識していた情報か	3. 0	2. 5	3. 55	0. 11
普遍的な情報か	3. 9	3. 8	0. 53	0. 49
製品設計に役立つか	3. 7	3. 1	8. 44	0. 03*
販売推進に役立つか	3. 5	3. 3	2. 63	0. 16

アンケートの回答は、5が肯定的、1が否定的。*は $p < 0.05$ を示す。

表3 アイデア抽出率

	C I 平均値	F G 平均値	F 値	P 値
観察データに対するアイデア抽出率	35. 7%	23. 8%	6. 25	0. 05*
分析結果に対するアイデア抽出率	14. 3%	14. 3%	0	1

*は $p < 0.05$ を示す。