

HI 設計チェックリストとその有用性の評価

加藤 秀一[†] 堀江 憲一^{††}
小川 克彦^{†††} 木村 重良^{††††}

コンピュータシステムの使いやすさや分かりやすさ（ユーザビリティ）を、誰もが客観的に手軽に評価できる HI（ヒューマンインタフェース）設計チェックリストを開発しその有用性を検証した。本チェックリストは、(1)専門家でなくとも使いこなせ、(2)信頼性・客観性の高い定量的な結果が得られ、さらに、(3)実際に、改善にまで直結できること、を主な目的に開発された。合計 133 のチェック項目により、画面（9 セクション、69 項目）と操作性（7 セクション、64 項目）がチェックできる。チェック結果は、先に開発した HI 設計ガイドラインへの適合度を表わす定量的な値で得られ、チェック項目やセクション間のバランス評価、他システムとの比較評価ができる。一方、本 HI 設計チェックリスト自身のユーザビリティを検証するため、4 名の被験者の、半年間にわたるチェックデータやチェック時間を分析した結果から、チェックリストの未経験者でも、インタフェースの不適合な箇所を、客観的かつ網羅的に、短時間で検出できることを確認した。また、HI 設計チェックリストで、運用/開発中のコンピュータシステムのユーザビリティを評価した結果から、HI 設計時に見落とされやすい（守られない）ガイドライン項目や、レビュー時に重点的にチェックすべき項目を抽出できた。これらのユーザビリティ評価を通して、HI 設計チェックリストの有用性が検証できたと考えている。

A Human Interface Design Checklist and Its Effectiveness

SUICHI KATO,[†] KENICHI HORIE,^{††} KATSUHIKO OGAWA^{†††}
and SHIGEYOSHI KIMURA^{††††}

A human interface design checklist, a tool by which anyone can easily and objectively evaluate the usability of his or her system, has been developed and its effectiveness has been evaluated. A total of 133 check items in two categories form this checklist. One is the category for screen check and includes 69 items classified into nine sections. The other is for dialogue check and includes 64 items classified into seven sections. Using the checklist, the degree to which the system conforms to our HI design guidelines can be estimated. This quantitative and objective measure represents the quality level of the screen design and dialogue control method and plays an important role when evaluating the balance between the sections in each category mentioned above and when comparing the quality level with those of other systems. The usability of the check method adopted in our checklist is also validated statistically. We also evaluated and improved 16 systems using this checklist. We gained analytical results showing that anyone, even a person with no experience with this checklist, can evaluate the quality level of his or her system. We also gained the check items with low conformity degree to which more attention should be paid.

1. ま え が き

コンピュータシステムのユーザ層の広がりや GUI（グラフィカル・ユーザ・インタフェース）の普及に

ともなって、システム開発者やソフトウェア設計者からは HI（ヒューマンインタフェース）ソフトウェアの設計ガイドラインが求められていた^{1)~4)}。一方、コンピュータユーザを中心に HI ソフトウェアの使いやすさや分かりやすさ（ユーザビリティ）を、誰もが客観的に手軽に評価できる HI 評価ツールが求められるようになってきた。

[†] NTT アドバンステクノロジー(株)
NTT Advanced Technology Corporation
^{††} NTT ヒューマンインタフェース研究所
NTT Human Interface Laboratories
^{†††} NTT サービス生産本部
NTT Service Engineering Headquarters
^{††††} NTT 東京技術開発センター
NTT Tokyo Technology and Development
Center

ユーザビリティの評価基準やモデル、ツールはいくつか示されている^{5)~7)}。開発当初から手軽に活用できるものに、ガイドラインやチェックリスト、解析のモデル、等が、開発後期のプロトタイプや製品を対象と

したものに、パフォーマンス測定やプロトコル分析、等がある。前者は比較的短時間で結果が得られる反面、結果の精度は粗い。後者は逆に、長時間を要するが得られる精度は高いといえる。実際には、開発段階や評価対象に応じてこれらを組み合わせながら評価を進めるのが望ましいとされている。

チェックリストは HI ソフトウェアの仕様段階から活用できるツールであり^{8),9)}、中でも 679 項目からなる Smith と Mosier のチェックリスト⁸⁾では、定量的な評価結果が得られる。チェック項目は、彼らが開発した 944 項目のガイドライン¹¹⁾に基づいており、チェックリストの利用者が判断する各チェック項目ごとの重要度 (Weighting) と、チェック項目に対応するガイドライン項目への適合度 (Rating) により、セクション別 (データ入力法、表示法、ユーザガイダンス、等) の定量的評価やシステム全体の評価ができる。ただし、チェックリストの利用者の立場 (システム開発者、ソフトウェア設計者、ユーザ) により重要度や適合度の見解が異なる恐れがあることから、客観性の保証までは検証されていない。また、実際の画面を見たり操作しながら網羅的にチェックする方法ではなく、チェックリストだけで評価するため、改善すべきセクションやチェック項目をチェックリスト上では指摘できても、実際の画面上で個別に指摘することは難しいと思われる。

著者らも、約 370 項目の HI 設計ガイドライン^{3),4)}に続き、HI 設計チェックリスト¹⁰⁾⁻¹²⁾によるコンピュータシステムのユーザビリティ評価手法を開発し、主に通信システムの開発者や設計者、ユーザに提供してきた。これと並行して、HI 設計チェックリスト自身の使いやすさや客観性の評価、ならびに、HI 設計チェックリストによる通信システムのユーザビリティ評価を進めてきた。

開発した HI 設計チェックリストは、133 のチェック項目からなり、画面 (9 セクション, 69 項目) と操作性 (7 セクション, 64 項目) がチェックできる。チェックは実際の画面を見たり操作をしながら網羅的に行い、チェック結果として、HI 設計ガイドラインへの適合度を表わす定量的な値が得られる。定量的な評価値により、チェック項目やセクション間のバランス評価、さらには、他のシステムとの比較評価ができるとともに、改善すべき箇所を実際の画面上で個々に指摘できる。一方、HI 設計チェックリスト自身のユーザビリティ評価実験を、半年間にわたって実施し

た結果、HI 設計チェックリストを初めて利用する場合でもインタフェースの不適合な箇所を、客観的かつ網羅的に、短時間で検出できることを確認した。また、HI 設計チェックリストによるコンピュータシステムのユーザビリティ評価結果から、HI 設計時に見落とされやすい (守られない) ガイドライン項目や、設計レビュー時に重点的にチェックすべき項目を抽出できた。これらのユーザビリティ評価を通して、HI 設計チェックリストの有用性が検証できたと考えている。なお、以下では HI 設計チェックリスト、HI 設計ガイドライン、チェック項目を、各々、チェックリスト、ガイドライン、項目と呼ぶことがある。

2. HI 設計チェックリスト

2.1 開発のねらい

ユーザビリティの高いチェックリスト、すなわち、

- (1) 専門家でなくとも使いこなせ、
- (2) 信頼性・客観性の高い定量的な結果が得られ、

さらに、

- (3) 実際に、改善にまで直結できること、
- をチェックリスト開発のねらいとした。これらを達成することで、(1)チェックリストの利用者の立場 (開発者、設計者、ユーザ) や利用経験による評価結果のばらつきがなくなる。また、(2)5段階評価のように主観が入りやすい結果ではなく、明確に規定された評価法に基づく定量的で客観的な結果が得られ、項目やセクション間のバランス評価や、他のシステムとの比較が可能となる。さらには、(3)評価だけに留まらず、HI の具体的な改善箇所や改善法に直接結びつき、開発期間やコストの削減、品質の保証が可能となる。

2.2 構成

多くの場合、コンピュータシステムの画面は、データの入力画面、データや表、メッセージの出力画面、メニューや属性の選択画面に分類でき、業務の流れに応じてこれらの画面が展開される。チェックリストは、これらの画面の展開に合わせて、画面と操作性が効率的にチェックできるように、画面チェックリスト (9 セクション, 69 項目) と操作チェックリスト (7 セクション, 64 項目) から構成されている。画面チェックリストの 9 セクションは、データ入力、データや表出力、等の画面の用途に応じたチェック項目、全画面に共通的なチェック項目、複数画面の展開法をチ

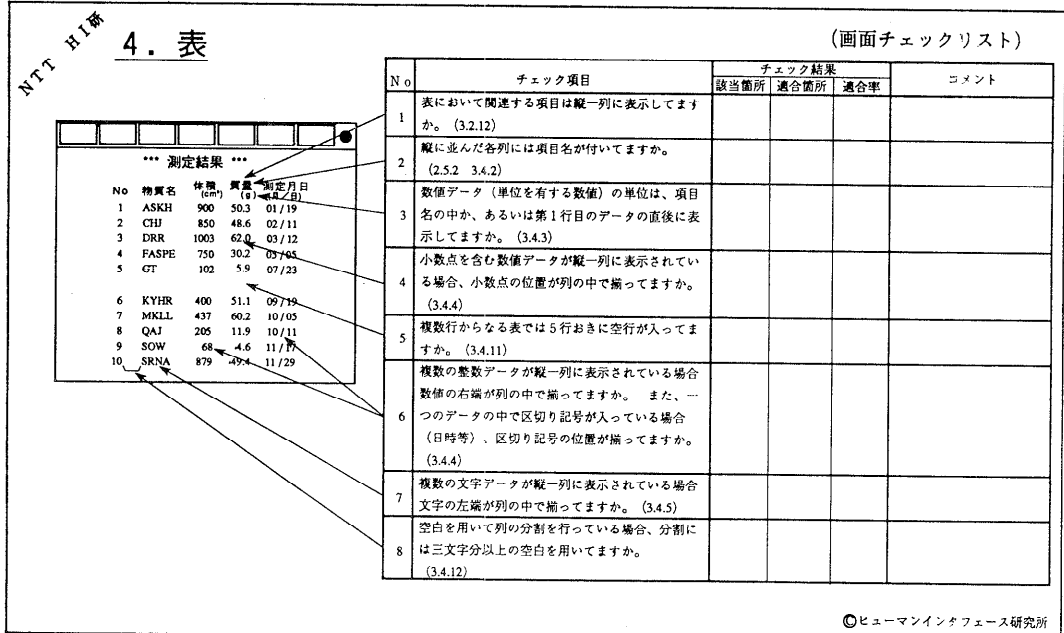


図 1 HI 設計チェックリストの構成 (表セクションの例)
Fig. 1 Human interface design checklist (table data check section).

表 1 HI 設計チェックリストの主なチェック内容
Table 1 Check items of human interface design checklist.

(a) 画面チェックリスト			(b) 操作チェックリスト		
セクション名	項目数	チェック内容	セクション名	項目数	チェック内容
データ入力	9	入力領域名、入力項目のグループ化、書式、プロンプト表示法	入力操作	17	質問形式、先頭機能、保護領域、確認、自動挿入、ゼロやスペースの自動挿入
データ出力	8	表示領域名、出力項目のグループ化、書式、単位	メニュー	4	初期メニューや直前メニューへの復帰、選択肢の数
メニュー	7	選択肢のグループ化、表示形式、属性のサンプル表示	コマンド	5	区切り記号、空白数、訂正法、連続起動
表	10	表示項目名、単位、分割法、見やすさ	エラー処理	6	繰り返しエラーへの対処法、メッセージの表示契機
メッセージ	9	肯定形や能動態の使用、操作順序との対応、色、エラー対処法	フィードバック	4	進行状況表示、完了表示、誤操作への対応
グラフ	11	注釈文、ラベル、単位、属性、強調表示法	応答時間	6	応答時間、エラーメッセージ表示契機、マイクロロール速度
流れ図	4	処理順序と流れの方向、強調表示法	共通項目	22	操作の確認、取消機能、中断/継続機能、ヘルプ、ポインティング領域の大きさ、操作モード表示、画面遷移
共通項目	6	画面名、色の数、番号付与基準、アイコン			
複数画面	5	画面の識別番号、対話やメニュー履歴、強調表示			

チェックする項目に分類されている。操作チェックリストも同様に、入力操作、応答時間、等のチェック項目、共通のチェック項目に分類されている。図 1 に、表セクションを例に画面チェックリストの構成を、表 1 に、各セクションの主なチェック内容を示す。なお、チェック範囲は、業務内容やシステム条件に依存しない共通的な範囲に限定している。

2.3 評価指標

Smith と Mosier のチェックリスト⁸⁾では、評価の最小単位をチェック項目とし、各項目ごとの重要度と適合度の積から、項目やセクション、システム全体の評価値を算出しており、重要度や適合度の判定に主観が入り込む余地がある。本チェックリストではより客観性を高めるため、評価の最小単位を、各項目でチェックされる個々のチェック箇所としている。その

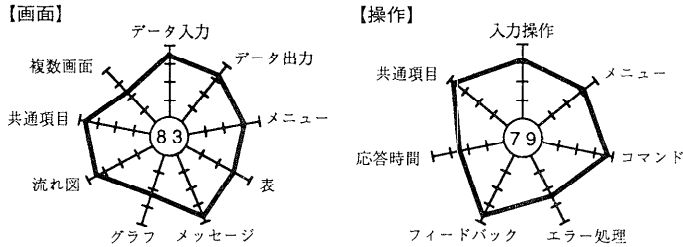


図2 評価結果の例 (適合指数とセクション別適合率)
Fig. 2 Example of evaluated measure (conformity degree).

ため、チェック対象とする全画面に対し、各項目ごとにチェックすべき箇所をすべて数え上げ (その数を項目別チェック箇所数と呼ぶ)、個々のチェック箇所ごとに適合度を Yes (適合) または No (不適合) のいずれかで判定し、項目別チェック箇所数に対する適合箇所数の割合 (項目別適合率と呼ぶ) を、チェック項目の評価指標としている。セクションの評価指標は、そのセクションに含まれる全項目の、項目別チェック箇所数の総和に対する、適合箇所数の総和の割合とし、セクション別適合率と呼んでいる。画面および操作全体の適合率も同様に定義し、これらは特に、画面または操作の適合指数と呼んでいる。従って、適合指数は全体的な品質を代表している値である。

本チェックリストでは、項目別チェック箇所数が多ければ多いほど、その項目は重要であると考え、それらを、セクション別適合率や適合指数算出時の重みとしている。チェック対象画面や操作の特性を反映した、よりダイナミックな評価手法と言える。Smith と Mosier の評価手法と比較すると、項目別チェック箇所数が重要度に、項目別適合率が適合度に対応する。

適合率や適合指数は、ガイドラインに適合していない項目の検出や、セクション間のバランス比較、システム間の客観的比較の指標となり、適合率が低い項目やセクションは重点的に改善すべき候補となる。例えば、セクション別適合率と適合指数が図2で得られたとすると (レーダーチャートの真ん中の数値が適合指数)、画面は全チェック箇所の 83% が適合しているものの、グラフと複数画面セクションが、また、操作性は 79% が適合しているものの、エラー処理と応答時間セクション、が重点的に改善すべき候補であることを示している。

2.4 評価と改善のステップ

チェック対象システムの知識を持ち合わせていない場合に、チェックリストで評価し改善案を作成するま

での典型的なステップを示す。システムの知識を持ち合わせている場合には、ステップ①と②は省略できる。

①チェック対象システムの調査

評価に当たっては、システム構成や業務内容、等の知識が必要であり、これらを事前に把握する。

②チェック対象画面の選定

代表的なタスク (例えば、メールの送信、障害申告の受付、障害履歴の検索タスク) で展開される画面を選定する。通常、連続した 5~6 画面で構成されることが多い。

③チェックリストによるチェック

選定した全画面を対象に、項目別チェック箇所をすべてチェックし、評価する。

④評価

項目別適合率、セクション別適合率、適合指数、レーダーチャート、等を分析し、チェック箇所数が多いにもかかわらず適合率が低い項目やセクションが存在すれば、それらは重点的に改善されるべきであることを指摘する。

⑤改善案の作成

不適合箇所をすべて改善した画面を作成する。改善法はチェック項目から容易に判断できる (チェック項目はガイドラインの裏返しとなっており、ガイドラインに立ち返らなくとも改善は容易である)。

⑥ドキュメントの作成

結果をドキュメントで提示する。評価結果だけでなく、具体的に不適合箇所や改善案までを提示すると、開発者や設計者の納得が得られやすい。

3. HI 設計チェックリストのユーザビリティ評価

チェックリストが、(1)専門家でなくとも使いこなせるか、(2)客観性の高い評価結果を保証できるか、を実験により検証した。実験では、画面チェックリストだけを使用した。操作チェックリストは機能の有無や操作の是非のチェックが中心であり、ほとんどの項目は項目別チェック箇所数が 1 となり、試行実験からも、チェック結果のばらつきは少ないと予測したためである。

3.1 評価実験手法

3.1.1 被験者

ガイドラインやチェックリストの利用経験を有する

経験者1名(男性, 経験2年)と, 未経験者3名(男性1名, 女性2名)を被験者とした。年齢はいずれも20代である。

3.1.2 タスク

未経験者3名には, 実験タスクに先立って学習タスク(2~3システム, システム当たり平均6画面)により, チェックリストの利用経験を積ませた(2~3週間)。

実験タスクでは, 通信関連の実システム(監視・制御, 故障受付・診断, 情報案内, 等のシステム)から, 代表的なタスクを選定し, 1システムごとに, 節2.4で述べた評価と改善の6ステップを実施し, 合計10システムを順に評価した。各被験者は, ステップ①, ②を共同で, ステップ③を個別に実施し, ステップ③終了時点で各自のチェック結果を持ち寄り, 合議の上で正しいチェック結果を判定した。経験者といえどもチェック箇所の見落としや数え間違い, 等があるためである。従って, 各被験者は1システムごとにチェック結果の正誤が分かる。ステップ④, ⑤, ⑥は各システムの評価責任者1名(いずれも未経験者が順に担当)だけが実施した。実験期間は約半年間である。

3.1.3 収集データ

各システムごとに, ステップ③で要した時間(被験者の申告値)と, 項目別正否, システム正解率を収集した。項目別正否は, その項目の項目別チェック箇所数と適合箇所数がともに正しい場合のみ正解とした。従って, 1箇所でもチェック箇所を見落としたり適合/不適合の判定を誤ると否となる。システム正解率は, 正しくチェックできた項目の割合を意味し, 項目別チェック箇所数が1以上の項目(適用された項目)を1項目以上含むセクション(適用されたセクション)だけを対象に求めた。すなわち, 適用されたセクションに含まれる項目数の総和に対する, 項目別正否が正解の項目の総数である。全項目とも適用されなかったセクションは, システム正解率の算出対象からは除外した。

実験タスクで評価した10システムの平均値は, チェック対象画面数が6画面, 総チェック箇所数(項目別チェック箇所数の総和)が約240箇所, 適用された項目数が約43である。これより, 適用された項目だけの項目別チェック箇所数は約5.6, 画面当りの総チェック箇所数は約40となる。

3.2 解析結果

3.2.1 正解率

経験を主要因, 反復測定した10種類のシステムを被験者内要因とした, 分散分析によるシステム正解率の有意差検定では, 経験の効果に有意な差があるとは言いきれない。一方, システムの効果には有意差があるが, 両者の交互作用に有意差はない。

経験	$F(1, 2) = 8.876, p = 0.097$
システム	$F(9, 18) = 16.925, p < 0.001$
交互作用	$F(9, 18) = 1.811, p = 0.136$

図3にシステム正解率の平均値を示す。システムによる差が大きいのは, 各システムの特徴の現われであり, 経験者といえども初めてチェックするような画面が含まれていたことによる。

3.2.2 チェック時間

分散分析によるチェック時間の有意差検定では, システム正解率同様, 経験の効果に有意な差は認められないが, システムの効果に有意差がある。両者の交互作用に有意差はない。

経験	$F(1, 2) = 5.969, p = 0.135$
システム	$F(9, 18) = 9.631, p < 0.001$

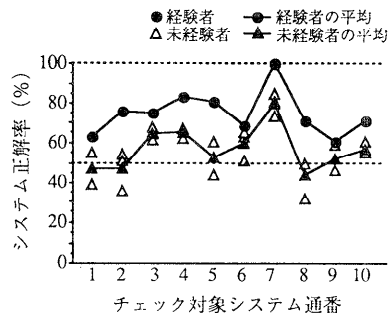


図3 平均システム正解率
Fig. 3 Mean rate of accuracy.

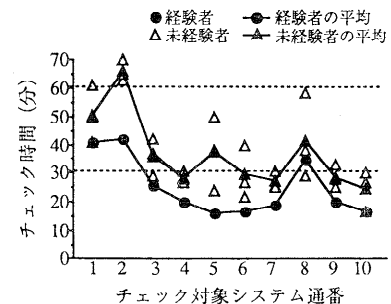


図4 平均チェック時間
Fig. 4 Mean time required for evaluation.

表 2 項目別チェック箇所数が多いチェック項目の適合率
Table 2 Check items with highest or lowest conformity degree.

セクション名	適合率が高いチェック項目	適合率が低いチェック項目
データ入力	データ入力領域には名前がついていますか。	データ領域名は、データ入力領域に近接していますか。
	データ領域名が入力領域の横に表示されている場合、領域名と入力領域の間には、一つ以上の空白（あるいは「:」等、空白に相当する区切り記号）がありますか。	入力データの文字数が指定されている場合、入力領域の長さを視覚的に案内（例えばアンダーライン等により）していますか。
データ出力	データ領域名は、そのデータの表示領域に近接していますか。	
	データ領域名が数値データや、文字データの横に表示されている場合、領域名とデータの間には一つ以上の空白（あるいは「:」等空白に相当する区切り記号）がありますか。	
	8文字以上の記号データや、数値データは、7文字以内のグループに分割していますか。ただし、7文字以内のデータについても習慣的に複数のグループに分けて取り扱っているデータ（例えば電話番号等）は、習慣に合ったグループに分けて表示していますか。	
表	表において関連する項目は縦一列に表示していますか。	縦に並んだ各列には項目名が付いていますか。
	複数の整数データが縦一列に表示されている場合数値の右端が列の中で揃っていますか。また、一つのデータの中で区切り記号が入っている場合（日時等）、区切り記号の位置が揃っていますか。	空白を用いて列の分割を行っている場合、分割には三文字分以上の空白を用いていますか。

交互作用 $F(9, 18) = 0.452, p = 0.888$

図 4 に平均チェック時間を示す。システムの効果の有意差は、システムにより画面数やチェック箇所数に差があるためであり、チェック時間の大小は総チェック箇所数にほぼ比例している。なお、平均チェック時間は、10 システム平均で、経験者が約 25 分、未経験者が平均約 37 分である。

4. HI 設計チェックリストによるコンピュータシステムのユーザビリティ評価

チェックリストによるコンピュータシステムのユーザビリティ評価結果をもとに、HI 設計時に見落とされやすいガイドライン項目や、設計レビュー時に重点的にチェックすべき項目を示す。

4.1 評価対象システムと評価指標

チェックリスト自身のユーザビリティ評価実験で評価したシステムに、通信量や通信網の監視・制御、等の実システムを加え、合計 16 システムのユーザビリティを画面チェックリストで評価し、項目別適合率やチェック箇所数、セクション別適合率、適合指数、を分析した。

4.2 評価結果

4.2.1 チェック箇所数

16 システムの平均は、画面当りの総チェック箇所数が約 43 箇所、画面当りの項目別チェック箇所数（項目別チェック箇所数/画面数）が、全チェック項目（69 項目）の平均で約 0.6 箇所、最大の項目で約 6.8

箇所、次いで、約 4.1 箇所であった。これらは、ともにデータ出力セクションのチェック項目“データ表示領域には名前がついていますか。”と、“データ表示領域名は、そのデータの表示領域に近接していますか。”である。一方、0.2 箇所未満のチェック項目も半数近くあった。

画面当りのセクション別チェック箇所数は、データ出力（15.3 箇所）、データ入力（9.3 箇所）、表（8.5 箇所）、の 3 セクションで全体（約 43 箇所）の約 75% を占めた。これらのセクションでチェックすべき箇所は、画面当たり平均 30 箇所以上と非常に多いことを示している。

4.2.2 適合度

一度も適用されなかった 13 の項目を除いた、56 項目の項目別適合率は、全システムを平均すると、約半数が 80% 以上を満足したが、70% 未満の項目もいくつかあった。表 2 は、項目別チェック箇所が多かった（画面当たり 1.0 箇所以上の）項目だけを対象に、適合率が高かった（80% 以上の）項目と低かった（70% 以下の）項目を示している。なお、項目別適合率と画面当りの項目別チェック箇所数との間に相関はみられなかった ($p = 0.530$)。

一方、セクション別適合率は、グラフ、共通項目、メッセージセクションで高く、いずれも 80% 以上であるが、複数画面、データ入力、表セクションは、他に比較するとやや低くなっている。データ入力や表セクションはチェック箇所数も多く、画面設計上で、特

に留意すべきセクションと言える。

適合指数に関しては、現実には、システムの制約(表示可能な文字数, 文字や行ピッチ, 文字フォントや色の制限, 等)があるため, 100%の実現は難しいが, 大部分のシステムは適合指数が75%以上であった。90%以上を達成しているシステムもいくつかあった。

5. 考 察

5.1 チェックリスト自身のユーザビリティ

チェックリストのユーザビリティ, すなわち, (i) 未経験者でもチェックリストを使いこなせるか, (ii) 客観的に評価できるか, (iii) 改善案の作成は容易か, を考察する。

第一の点に関し, システム正解率の実験結果では, チェックリストの利用経験による差は認められない。未経験者のシステム正解率は51%~61% (平均57%)と, 経験者(平均75%)に比較して低いように思えるが, 項目別正否を, その項目の項目別チェック箇所数と適合箇所数とともに正しい場合のみ正解とし, 1箇所でもチェック箇所を見落としたり適合/不適合の判定を誤ると否としたためであり, 未経験者でも十分にチェックリストを使いこなす, ほとんどの不適合箇所を検出できている。また, チェック時間にも経験差は認められないことから, 未経験者でも手軽に(画面当たり5~6分のチェック時間で)チェックリストを利用できると期待している。

なお, より高精度の評価結果を得るためには, (i) 何人かでチェックする, (ii) 適合率や適合指数の算出単位を小さく(例えば, 画面単位に)し, それらの結果からシステムの評価指標を算出する, 等が効果的と考えられる。

第二点に関しては, チェック箇所の数え上げ方やYes/Noの判定基準が明確であるか否かが問われるが, チェックした7000以上の箇所はすべてYes/Noで評価できていること, システム正解率やチェック時間に, 経験による有意差がないことから, 客観性の高い定量的評価手法が提供できたと考えている。

一方, チェックリストで満足できる結果を得ることは, 必要条件ではあるが十分条件ではない。チェックリストでは, 初心者にもやさしいインタフェースが実現されているか否かを, いくつかの観点からチェックしている。単純そうな画面でも20~30箇所, 複雑な画面では100箇所以上をチェックし, すべての不適合

箇所(著者が評価した16システムの平均では, 全チェック箇所数の約20%, 数にして画面当たり約10箇所)を指摘できるが, 最終的なHI設計には, 実ユーザによる評価実験結果, 等をも考慮する必要がある。

最後の点に関して, チェックリストの各項目には, 根拠となったガイドラインとの対応関係が明記されていること, また, 実験の被験者白らが設計した改善案は, 開発者や設計者, ユーザに問題なく受け入れられていることから実証できたと考えており, 設計レビュー時に使用する効果は大きいと言える。

なお, 第一の点の, 未経験者でもチェックリストを使いこなせるかに関してであるが, システム正解率の実験結果からは, チェックリストの利用経験による差は認められないが, どのシステムの正解率とも, 経験者が未経験者より高いことから, 今後とも実験を継続していく予定である。

5.2 コンピュータシステムのユーザビリティ

コンピュータシステムのユーザビリティ評価結果をもとに, HI設計時に見落とされやすいガイドライン項目や, 設計レビュー時に重点的にチェックすべき項目を抽出した(表2)。適合率を低くしている主な要因は, 視覚的な対応やガイダンスの不備, 字数や行数が限られている画面に多くの(あるいは, 不必要とも思われる)情報を一度に表示したために項目名や列分割スペースを省略した, 等である。このほかの要因としては, データやメニューのグループ化, 表データの表示形式(5行おきに空行を入れる), 画面展開(画面のタイトル, ページ識別子)の不備があげられる。これらは, 主に通信システムの分析から得られた結果であるが, 通信システム以外のコンピュータシステムにも通じる点が多いと思われ, HI設計やレビュー時に重点的にチェックすべき項目と言える。

6. む す び

コンピュータシステムのユーザビリティ(使いやすさや分かりやすさ)を, 誰もが客観的に手軽に評価できるHI設計チェックリストを開発し, その有用性を評価した。HI設計チェックリストは, 133のチェック項目からなり, 画面(9セクション, 69項目)と操作性(7セクション, 64項目)がチェックできる。チェック結果は, 先に開発したHI設計ガイドラインへの適合度を表わす定量的な値で得られ, チェック項目やセクション間のバランス評価, 他システムとの比較評価ができる。

一方、HI 設計チェックリスト自身のユーザビリティは高く、専門家でなくとも、インタフェースの不適合な箇所（全チェック箇所数の約 20%、数にして画面当たり約 10 箇所）を、網羅的に、短時間で容易に検出でき、不適合箇所の改善案の作成も容易である。さらに、チェックリストの副次的効果として、設計ガイドラインやチェックリストの背景となっている設計ノウハウが蓄積され、改善案の作成や設計能力の向上が見られた。また、コンピュータシステムのユーザビリティ評価結果からは、HI 設計時に見落とされやすいガイドライン項目や、レビュー時に重点的にチェックすべき項目が抽出できた。これら両者により、HI 設計チェックリストの有用性が検証できたと考えている。今後、チェック範囲を広げるとともに、システム構成やサービス特性、ユーザ特性（システムを使用するユーザの年齢層、経験度、使用頻度、等）をも反映できる評価体系へと発展させたい。また、システム正解率の評価実験を継続していく。さらに、ISO で進められている HCI の国際標準化⁹⁾の進展にあわせ、それらとも整合をとって行きたい。

謝辞 本研究の機会を与えていただいたヒューマンインタフェース研究所 高野所長、釜江前所長、マルチメディア処理研究部 栗原部長、遠藤前部長に深謝いたします。また、評価にあたり非常に多くの方にご協力いただきました。心よりお礼申し上げます。

参考文献

- 1) Smith, S.L. and Mosier, J.N.: Guidelines for Designing User Interface Software, Technical Report ESD-TR-86-278, MITRE, MA (1986).
- 2) Shneiderman, B.: *Designing the User Interface*, Addison-Wesley Publishing Co. Inc. (1987). (東, 井関 (監訳): ユーザ・インタフェースの設計, 日経マグローヒル社 (1987).)
- 3) 米村, 小川: 通信システムにおけるヒューマンインタフェース設計指針の考察, NTT R & D, Vol. 39, No. 2, pp. 249-256 (1990).
- 4) 米村, 小川: ブックメタファを用いた HI 設計ガイドラインデータベース, 情報処理学会, 第 36 回ヒューマンインタフェース研究会, 36-2 (1991).
- 5) Whiteside, J., Bennett, J. and Holtzblatt, K.: Usability Engineering: Our Experience and Evolution, Helander, M. (ed.), *Handbook of Human-Computer Interaction*, pp. 791-817 (1988).
- 6) 浜田, 森川, 小川: HCI の国際標準化~その 1: 標準化の視点とユーザビリティ~, 情報処理学会, 第 45 回ヒューマンインタフェース研究会, 45-14 (1992).

- 7) 旭, 岡田, 井関: 「使いやすさ」の計測: UI テスタ, 計測自動制御学会 HI 部会, 第 8 回ヒューマン・インタフェース・シンポジウム論文集, pp. 287-290 (1992).
- 8) Smith, S.L. and Mosier, J.N.: A Design Evaluation Checklist for User-System Interface Software, Technical Report ESD-TR-84-358, MITRE, MA (1984).
- 9) Ravden, S. and Johnson, G.: *Evaluating Usability of Human-Computer Interfaces*, Ellis Horwood Limited (1989). (東 (監訳): ユーザ・インタフェースの実践の評価法—チェックリストアプローチによる使いやすさの向上—, 海文堂 (1993).)
- 10) 木村, 米村, 小川: ソフトウェアのユーザビリティチェックリスト, 日本人間工学会, 第 21 回関東支部大会, A 1-16, pp. 63-64 (1991).
- 11) 加藤, 堀江, 小川, 木村: HI 設計チェックリストとそのユーザビリティ評価, 情報処理学会, 第 50 回ヒューマンインタフェース研究会, 50-3 (1993).
- 12) 堀江, 加藤, 小川, 木村: HI 設計チェックリストとその有用性, 計測自動制御学会 HI 部会, 第 9 回ヒューマン・インタフェース・シンポジウム論文集, 1523, pp. 321-324 (1993).

(平成 6 年 2 月 2 日受付)

(平成 6 年 10 月 13 日採録)



加藤 秀一 (正会員)

昭和 46 年名古屋大学工学部電気工学科卒業。昭和 48 年同大学院工学研究科修士課程修了。同年日本電信電話公社 (現 NTT) に入社。計算機複合体の研究, 通信処理プログラム, 分散処理 OS の実用化, リアルタイム AI, ヒューマンインタフェースの研究に従事。平成 6 年 2 月 NTT アドバンステクノロジー社に出向。HIT センタにおいて HI ビジネスに従事。電子情報通信学会, 人工知能学会, HFES 各会員。



堀江 憲一

NTT ヒューマンインタフェース研究所映像処理研究部勤務。昭和 63 年入社。平成 4 年よりヒューマンインタフェースの設計・評価業務に従事。



小川 克彦 (正会員)

NTT サービス生産企画部担当部長。昭和 53 年入社。画像情報システムの実用化, ISDN の技術企画, ヒューマンインタフェースの研究に従事。昭和 51 年慶應義塾大学工学

部管理工学科卒業。昭和 53 年同大学院工学研究科修士課程修了。工学博士。電子情報通信学会, 人間工学会, IEEE, HFES 各会員。昭和 60 年電子情報通信学会学術奨励賞受賞。



木村 重良

NTT 東京技術開発センタ勤務。昭和 58 年入社。平成 2 年より 5 年までヒューマンインタフェース研究所にてヒューマンインタフェース設計・評価業務に従事。