

作業時間の見積もり能力と非認知能力

竹内俊彦^{*1 *4}, 若山昇^{*2 *4}, 立野貴之^{*3}, 山本美紀^{*2 *4}, 草山太一^{*2}

^{*1} 駿河台大学 ^{*2} 帝京大学 ^{*3} 松蔭大学 ^{*4} 教育テスト研究センター

Ability in estimate working hours of a task and Non-Cognitive Skill ability

Toshihiko Takeuchi^{*1 *4}, Noboru Wakayama^{*2 *4}, Takashi Tachino^{*3},
Miki Yamamoto^{*2 *4}, Taichi Kusayama^{*2}

^{*1} Surugadai University, ^{*2} Teikyo University, ^{*3} Shoin University,
^{*4} Center for Research on Educational Testing

We conducted an experiment on 71 students to estimate the time it would take to solve simple calculations. We then investigated the correlation between their ability to estimate, critical thinking scores and their answers on a questionnaire regarding their personalities. We found a significant positive correlation between reasoning power, tenacity, and empathy, as well as a significant negative correlation with weaker intentions, slower work, and lower self-motivation. It was also found that the degree of optimism in the estimates was strongly correlated with tenacity, empathy, self-confidence, and willpower, and negatively correlated with weakness, lack of self-confidence.

キーワード: 非認知能力, 見積もり能力, 測定, 評価, アンケート

1. はじめに

我々は非認知能力を評価する方法を研究している。非認知的能力とは従来のテストでは測りにくいが社会で生きていくために必要な力のことである(池迫, 宮本 2015)。たとえば「目標に向かって頑張る力, 他の人とうまく関わる力, 感情をコントロールする力」などがその代表とされている。

非認知能力を計測する手法は確立されていない。定義からしても非認知能力を従来型の筆記テストで測定するのは難しいと予想される。

そこで我々は, 非認知能力を測定する手法を過去に提案した。たとえば短時間に多数回のクリックをしてお神輿を担ぐ単純なゲームをさせる方法を提案し(竹内ら 2020), その結果とクリティカル・シンキング能力など他のテストとの相関を調べ分析を行った(山本ら 2020)。また, 同じシチュエーションを2つの立場から描いたマンガについて, それぞれ自分だったら相

手の行動をどう評価するかを多数の人に回答していただく方法(竹内ら 2020)も発表している。これは多数のデータから, 考え方の偏りや, どんな場合にトラブルがどれくらいの確率で起きやすいかを回答者に示すことで「他の人とうまく関わる力」を測る試みである。

一連の研究が最終的に目指すのは「受験者を現実を模したシミュレーション世界に置き, その世界での行動を記録する」「多数の受験者に行ってデータを蓄え, その集合知から非認知的能力を計る。つまり, 望ましい行動を事前に仮定するのではなく, 集合知の結果によって望ましいとされる行動を決め, シミュレーションである受験者の他者からの評価を予測する」ことである。

今回は研究のスタートとして, ある作業の完成までにかかる時間を事前に見積もらせ, その後に実際に作業してもらい, 見積もりの正確さを測定し, その値と他の能力との相関を調べた。

非認知的能力のうち、特に作業時間の見積もり能力に注目したのは、エビデンス・ベース・スケジューリングというシミュレーション方法と合わせると、会社に入ったときの仕事遂行の局面を、非常に抽象化された形ではあるが、コンピュータ・シミュレーションできると考えたからである。エビデンス・ベース・スケジューリングについては次章で述べる。

2. エビデンス・ベース・スケジューリング

エビデンス・ベース・スケジューリング(以下 EBS と記す)は Joel Spolsky がその著書「More Joel On Software」20 章で提案した、作業時間を見積もる手法である(Joel Spolsky 2009)。概要は以下である。

- [1] ある大きな仕事 P を複数の工程 $p_1 \sim p_n$ に分け、また、その仕事の完成に必要な時間 $pe_1 \sim pe_n$ を見積もらせる。
- [2] $p_1 \sim p_n$ の各工程それぞれに似た仕事 $q_1 \sim q_n$ を与え、完了に必要な時間 $qe_1 \sim qe_n$ を予想させた後で、実際に $q_1 \sim q_n$ を行わせ、かかった時間 $qr_1 \sim qr_n$ を測定する。
- [3] qr_k/qe_k の比率 $r_1 \sim r_n$ を計算する。つまり、予想の何倍だったかの比率を計算する。
- [4] その人の仕事における見積もり時間 $pe_1 \sim pe_n$ に、 $r_1 \sim r_n$ が等確率 $1/n$ で現れる乱数を掛け、その総和を取って 1 回のシミュレーションとする。そのシミュレーションを 100 回行って、その人が仕事を終了させる確率の度数分布表を得る。

EBS の概要を図 1, EBS で得られる度数分布表の例を図 2 に示す。

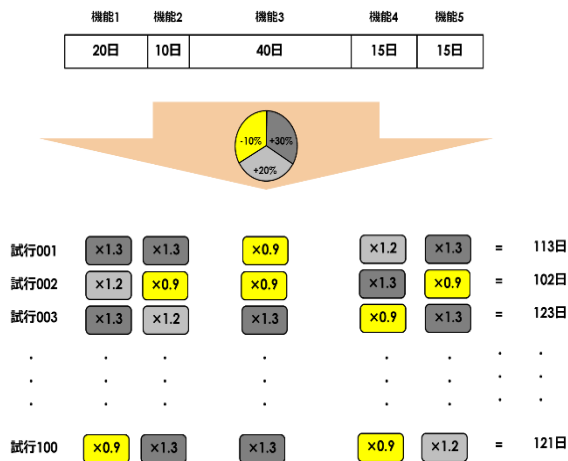


図 1 EBS の概要

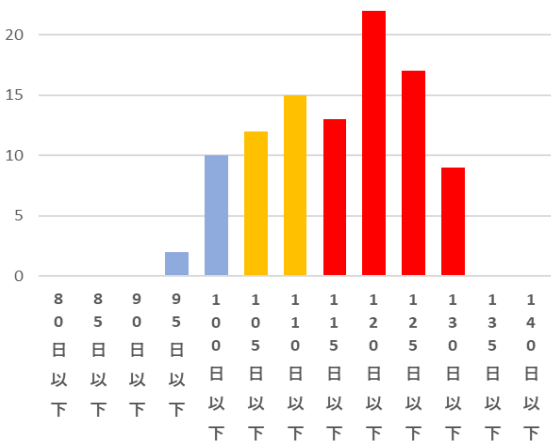


図 2 EBS で得られる度数分布表の例

将来的な構想では、各人の見積もり能力のデータを基礎にして、会社の業務遂行をシミュレーションし、そのシミュレーション内で各人がどう行動するかを記録し分析する。

たとえば上司から割り振られた仕事の締め切りが 1 週間後に迫っているのに進捗が思わしくないときどう行動するか、締め切りに間に合わないときどう解決するかなど、主要な場面を想定し回答することで、ある人がどんな困難に陥りやすく、それをどのように切り抜けるかのデータを得る。また立場を逆にして、自分が上司なら、部下が締め切りを破ったときどのように評価するか、といったデータにも回答してもらう。それらのデータを多数、集めてデータベース化する。すると、ある人の行動について、架空の会社における確率的なシミュレーションを多数回行うことができ、他人からどのように評価されるかがわかる。多数回のシミュレーション結果から、組織内で行動する能力、つまり非認知能力の測定ができるようになる、というのが我々の目指すところである。

3. 実験

作業時間の見積もり正確さ能力を測定するためのサイトを開発し、実験を行った。実験は 2020 年 11 月 22 日の 15:10-16:10 に、実験参加者を日本全国の各大学から募り、オンライン(Zoom)で行った。実験参加者は 75 名であったが、中途脱落者のデータを除くと、有効なデータは 71 名だった。分析に利用した実験参加者 71 名の属性を表 1 に示す。

表 1 実験参加者の属性

大学	45校
性別	男性 34人 女性 37人
文理	文系 43人 理系 28人
学年	1年 11人 2年 16人 3年 23人 4年 21人

なお本実験は、他の研究者との共同実験の一部として実施した。つまり実験参加者に謝金を払って 1 日、拘束し、9:00-17:30 まで、約 1 時間×6 種類の実験を行ったうちの 5 番目の実験である。ゆえに実験参加者は疲れていた可能性がある。また 5 番目の実験では仕事の見積もり正確度だけを測定したが、本研究の分析では、クリティカル・シンキングの能力や自分の性格に関するアンケートとの相関を示している。クリティカル・シンキングの能力や自分の性格に関するアンケートのデータは 4 番目の実験で取得したものである。

3.1 実験手順

実験の手順を以下に示す。

手順 1. 実験用に作ったサイトにアクセスし、実験 ID、性別、所属大学、学年を入力する

手順 2. 1 桁+1 桁の足し算を 10 問、22 秒で解く動画を見せた後に、実験参加者は 30 問を何秒で解けそうか予測させる。図 3 にステップ 02 の画面を示す。

ステップ02 事前・所要時間予想アンケート

Q01 あなたの実験IDは？
D60
*実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのでない限り、変えないでください。

次の22秒の動画を見てから、Q02に回答してください。

Q02 この動画のスピードなら、10問で22秒、かかっていますので、30問を解くなら約66秒となるでしょう。さて、あなたが同じような足し算を 30問 解くなら、何秒かかると予測しますか？
秒

それでは↓のボタンを押して、実際に10問、解いてみましょう。
10問の計算練習スタート

図 3 ステップ 02 の画面

手順 3. 練習として足し算を 10 問、実際に練習させる。なお正解を入力しない限り、次の問題に進まない。

図 4 にステップ 03 の画面を示す。

ステップ03 【練習】足し算問題

Q01 あなたの実験IDは？
D60
*実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのでない限り、変えないでください。

正解で次問へ。

問04 / 全10問

5+2 =

図 4 ステップ 03 の画面

手順 4. 実際に 10 問を解く練習をした後に、再度、30 問を解くとした時の予想時間を入力する。ステップ 04 の予想時間入力画面を図 5 に示す。

ステップ04 事後・所要時間予想アンケート

Q01 あなたの実験IDは？
D60
*実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのでない限り、変えないでください。

Q02 練習、お疲れ様でした。さきほどは10問を解いて、結果は 14秒でした。次はいよいよ本番です。今、体験した足し算を 30問 解くなら、今度は何秒かかると予測しますか？
問題数が3倍になるので、単純に時間も3倍かかると考えるなら、42秒となります。
秒

それでは↓のボタンを押して、実際に30問、解いてみましょう。
30問の計算 スタート

図 5 ステップ 04 の画面

手順 5.1 桁+1 桁の足し算を 30 問、実行する。画面を図 6 に示す。

ステップ05 【本番】足し算問題

Q01 あなたの実験IDは？
D60
*実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのでない限り、変えないでください。

正解で次問へ。

問04 / 全30問

8+8 =

図 6

ステップ 05 の画面

手順 6.2 桁+2 桁の問題を 30 問、解くときの時間を予想してもらう。ステップ 06 の画面を図 7 に示す。

ステップ07 【本番】足し算問題

Q01 あなたの実験IDは？
D60
*実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのでない限り、変えないでください。

正解で次問へ。

問03 / 全30問

13+22 =

図 7 ステップ 06 の画面

手順 7. 2 桁+2 桁の足し算を 30 問、実行する。画面を図 8 に示す。

ステップ06 事後・所要時間予想アンケート

Q01 あなたの実験IDは？

D60

！実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのではない限り、変えないでください。

Q02 さきほどは30問の足し算を解いて、結果は 35秒でした。

次は2桁と2桁の足し算です。2桁と2桁の足し算を 30問 解くなら、今度は何秒かかると予測しますか？

だいたい3倍くらいかな、と考えるなら、106秒となります。秒

それでは↓のボタンを押して、実際に30問、解いてみましょう。

30問の計算 スタート

図 8 ステップ 07 の画面

手順 8. 2 桁×1 桁の足し算を 30 問、30 問を解くとした時の予想時間を入力する。画面を図 9 に示す。

ステップ08 事後・所要時間予想アンケート

Q01 あなたの実験IDは？

D60

！実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのではない限り、変えないでください。

Q02 さきほどは30問の足し算を解いて、結果は 82秒でした。

次は2桁と1桁の掛け算です。は2桁と1桁の掛け算を 30問 解くなら、今度は何秒かかると予測しますか？

だいたい2倍くらいかな、と考えるなら、164秒となります。秒

それでは↓のボタンを押して、実際に30問、解いてみましょう。

30問の計算 スタート

図 9 ステップ 08 の画面

手順 9. 2 桁×1 桁の足し算を 30 問、実行する。画面を図 10 に示す。

ステップ09 【本番】足し算問題

Q01 あなたの実験IDは？

D60

！実験ID欄は、うっかり途中でブラウザを閉じてしまい、このステップを、このページからやり直すのではない限り、変えないでください。

正解で次問へ。

問01 / 全30問

12×3 =

図 10 ステップ 09 の画面

4. 分析結果

本実験では「1 桁+1 桁の足し算」「2 桁+2 桁の足し算」「2 桁×1 桁の掛け算」を、それぞれ 30 問ずつ解くときの見積もりを各実験参加者に行っていたいただいた後、実際に作業をしていただいた。

4.1 クリティカル・シンキングのテスト内容

本実験では、作業完了に必要な時間の見積もり能力

と、認知的能力の代表としてクリティカル・シンキングの能力の相関を調べた。

クリティカル・シンキング能力を測定する問題には、前の時間の実験で 15 問、出題し、回答いただいたデータをを用いた。分析に関する問題、推論に関する問題 5 問、読解に関する問題を 5 問ずつである。

4.2 アンケート内容

本実験で参照したアンケートの質問項目は全 68 問である。各質問文にどの程度同意できるかを、全くあてはまらない 1～非常にあてはまる 5 の 5 点法で質問した。質問項目は、目標に向かって頑張る力(「18. 数年にわたる努力を要する目標を達成したことがある。」等)、他の人とうまく関わる力(「31. 人が頑張っているのを見たり聞いたりすると、自分には関係なくても応援したくなる。」等)、感情をコントロールする力(「26. 私には、自分の感情を適切にコントロールする力が、とてもある。」等)である。詳細については先行研究(山本ら 2020)で述べたので割愛する。

4.3 見積もりの正確度

見積もりの正確度は「実際にかかった時間/見積もった時間・1」の絶対値とした。つまり、見積もった時間が 100 秒ならば、実際にかかった時間が 120 秒でも 80 秒でも、どちらも 0.2 となる。「1 桁+1 桁」「2 桁+2 桁」「2 桁×1 桁」という 3 種類の見積もりの正確度が得られるが、その平均値を 71 名それぞれについて計算した。3 種類の見積もりの正確度の平均値を、4.2 で説明したクリティカル・シンキングのテスト得点、4.3 で説明した全 68 問のアンケートと相関を調べた。

4.4 作業時間の予測の正確さとの相関

作業時間の予測の正確さと、クリティカル・シンキング能力、アンケートとの相関が高い項目を調べた。表 2 に相関が高かった(p 値が 10%以下の)項目を示す。

作業時間の見積もり能力と正の相関が高かったのは、クリティカル・シンキング能力のうち特に推論力、ねばり強さ、共感力といった項目だった。また負の相関が高かったのは、自分への満足度、意志の弱さ、30 問の問題を解く所要時間だった。全般的に、作業時間の見積もり能力は、他の「望ましい」と思われる性質と

相関が高い。作業時間の見積もり能力は、非認知能力の良い指標の候補となりそうである。

表 2 作業時間の予測の正確さと相関の高い項目			
「作業時間の予測の正確さ」との相関係数		意味	p値(%)
正の相関	Q06～Q10合計 クリティカル・シンキング推論力	0.286 推論力がある	1.5%
	Q08(推論力の問題)	0.278 推論力がある	1.8%
	51. 始めたことは、どんなことでも最後までやりとげる。	0.273 ねばり強い	2.0%
	31. 人が頑張っているのを見たり聞いたりすると、自分には関係なくても応援したくなる。	0.238 共感力がある	4.4%
	20. 悲しんでいる人を見ると、なぐさめてあげたくなる。	0.225 共感力がある	5.7%
	クリティカル・シンキング総合点	0.216 賢い	6.9%
	62. 誘惑に負けない。	0.213 ねばり強い	7.3%
	Q01～Q05合計 クリティカル・シンキング分析力	0.204 分析力がある	8.5%
	10. 趣味や娯楽のせいで、やるべきことがそっちのけになることがある。	-0.205 意思が弱い	8.4%
	Q12 所要時間	-0.212 仕事が遅い	7.4%
負の相関	55. 目標を決めても、後から変えてしまうことがよくある。	-0.245 意思が弱い	3.8%
	32. だいたいにおいて、自分に満足している。	-0.270 自分に甘い	2.2%

4.5 見積もりの楽観度

見積もりの楽観度は「実際にかかった時間/見積もった時間・1」とした。つまり、見積もった時間が100秒ならば、実際にかかった時間が120秒なら+0.2、80秒なら-0.2となる。「1桁+1桁」「2桁+2桁」「2桁×1桁」という3種類の見積もりの楽観度が得られるので、その平均値を71名それぞれについて計算した。3種類の見積もりの楽観度の平均値を4.2で説明したクリティカル・シンキングのテスト内容得点、4.3で説明した全68問のアンケートとの関係を調べた。

4.6 楽観度の分布

71名の楽観度の分布は図11のようになった。なお最大値は1.307(見積もり時間の1.307倍かかる)、最小値は0.662(見積もり時間の0.662倍で済む)、平均値は0.910であった。つまり平均的には、実際にかかった時間より1割ほど多めに見積もっていた。本当にそう思ったのか、真に「この時間でできる」と思った値に少し余裕を持たせて回答したかは不明である。

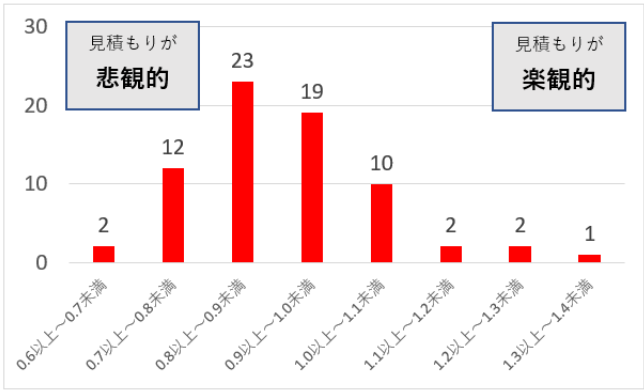


図 11 楽観度の分布

楽観度に性差、学年と関係があるか調べたが、強い関係はなかった。文系/理系については、カイ事情検定でp値=0.087とやや低かった。つまり理系の方が自分の計算速度について正確な見積もりを出す傾向があったのかもしれない。楽観度と文系/理系のクロス集計表を表3に示す。

表 3 楽観度と文系/理系のクロス集計表

楽観度	文系	理系	総計
楽観的 (0.95以上)	11	4	15
普通 (0.8以上0.95未満)	21	21	42
悲観的 (0.8未満)	11	3	14
総計	43	28	71

楽観度と性別のクロス集計表を表3に示す。

表 4 楽観度と性別のクロス集計表

楽観度	男性	女性	総計
楽観的 (0.95以上)	7	8	15
普通 (0.8以上0.95未満)	19	23	42
悲観的 (0.8未満)	8	6	14
総計	34	37	71

楽観度と学年のクロス集計表を表5に示す。

表 5 楽観度と学年のクロス集計表

楽観度	1年	2年	3年	4年	総計
楽観的 (0.95以上)	2	4	5	4	15
普通 (0.8以上0.95未満)	5	9	13	15	42
悲観的 (0.8未満)	4	3	5	2	14
総計	11	16	23	21	71

4.7 作業時間の予測の楽観度との相関

作業時間の予測の楽観度と、クリティカル・シンキング能力、アンケートとの相関が高い項目を調べた。表6に相関が高かった(p値が10%以下)項目を示す。

表 6 作業時間の予測の楽観度と相関の高い項目

	「作業時間の予測の正確さ」との相関係数	意味	p値(%)
正の相関	18. 数年にわたる努力を要する目標を達成したことがある。	0.287 ねばり強い	1.5%
	31. 人が頑張っているのを見たり聞いたりすると、自分には関係なくても応援したくなる。	0.270 共感力がある	2.2%
	30. いろいろな良い素質を持っている。	0.267 自信がある	2.4%
	35. 私は頑張る屋だ。	0.263 ねばり強い	2.5%
	62. 誘惑に負けない。	0.255 意思が強い	3.1%
	64. 物事を人並みには、うまくやれる。	0.227 自信がある	5.5%
	53. 少なくとも人並みには、価値のある人間である。	0.214 自信がある	7.1%
負の相関	60. 他人の成功を見聞きしているうちに、焦りを感じることが多い。	-0.214 自信がない	7.2%
	63. 新しいアイデアや計画によって、それまで取り組んでいたことから注意がそれることがある。	-0.220 意思が弱い	6.3%
	21. 自分には、自慢できることがあまりない。	-0.224 自信がない	5.8%
	Q06所要時間	-0.225 仕事が遅い	5.8%
	45. だらけてしまう。	-0.228 意思が弱い	5.4%
	47. 相手を批判するときは、相手の立場を考えることができない。	-0.230 共感力がない	5.2%
	Q07所要時間	-0.231 仕事が遅い	5.1%
	56. 何かにつけて、自分は役に立たない人間だと思う。	-0.232 自信がない	5.0%
	34. あるアイデアや計画に一時的に夢中になっても、あとで興味を失うことがある。	-0.234 意思が弱い	4.8%
	27. 自分は全くだめな人間だと思うことがある。	-0.237 自信がない	4.5%
	15. 自分にとってよくない誘いは、断る。	-0.242 意思が強い	4.1%
	7. 数ヶ月以上かかるような計画に集中して取り組み続けることは難しい。	-0.251 意思が弱い	3.3%

作業時間の予測の楽観度と正の相関が高かったのは、「数年にわたる努力を要する目標を達成したことがある.」「人が頑張っているのを見たり聞いたりすると、自分には関係なくても応援したくなる.」等、ねばり強さ、共感力、自信、意思の固さなどであった。また負の相関が高かったのは、「自分にとってよくない誘いは、断る.」という例外を除いては、「数ヶ月以上かかるような計画に集中して取り組み続けることは難しい.」「自分は全くだめな人間だと思うことがある.」など、おおむね意志の弱さ、自信のなさ、仕事の遅さ、共感力のなさを意味する項目だった。

5. おわりに

単純な計算問題を解く時間を、事前に見積もらせる実験を 71 名の大学生に対して行った。その後、見積もる能力と、クリティカル・シンキングの得点や自分の性格に関するアンケートとの相関を調べた。その結果、それほど相関係数は高くないが、推論力、ねばり強さ、共感力とは有意な正の相関が、意思の弱さ、仕事が遅い、自分に甘い、と有意な負の相関があるとわかった。また見積もりの楽観度とは、ねばり強さ、共感力、自信、意思の固さと有意な正の相関が、意志の弱さ、自信のなさとは有意な負の相関があることがわかった。仕事時間の見積もり能力は、非認知能力を EBS でシミュレーションし測定するときに、基礎データの候補になると結論した。

謝辞

本研究は科研費（課題番号 19K03013）と教育テスト研究センター(CRET)の助成を得た。

参 考 文 献

- (1) 池迫浩子, 宮本晃司: “家庭, 学校, 地域社会における社会情動的スキルの育成”, ベネッセ教育総合研究所, pp.40-50 (2015)
- (2) 竹内俊彦, 草山太一, 立野貴之, 山本美紀, 若山昇: “大学生の非認知的能力の計測を試みるゲーム開発”, JSiSE 第 45 回全国大会予稿集, pp.3-4 (2020)
- (3) 山本美紀, 草山太一, 竹内俊彦, 立野貴之, 若山昇: “非認知的能力に関する計測データの分析”, JSET 2020 年秋季全国大会予稿集, pp.173-174 (2020)
- (4) 竹内俊彦, 若山 昇, 草山太一, 山本美紀, 立野貴之: “人間関係トラブル予測能力を測定する方法の提案”, 日本教育メディア学会研究会論集 49, pp.68-72 (2020)
- (5) Joel Spolsky: “Joel On Software”, pp.155-166, 翔泳社, pp.68-72 (2009)