

感情分析を用いた Twitter の投稿内容による選手採点システムの提案

若狭 孟[†] 横山 昌平^{††}

[†] 首都大学東京 システムデザイン学部 〒 191-0065 東京都日野市旭が丘 6-6

^{††} 首都大学東京大学院 システムデザイン研究科 〒 191-0065 東京都日野市旭が丘 6-6

E-mail: [†]wakasa-takeshi@ed.tmu.ac.jp, ^{††}shohei@tmu.ac.jp

あらまし 近年, Twitter などの SNS が普及したことにより, その時々起きたことや感じたことをほぼリアルタイムに投稿する人が多く存在している. サッカーなどのスポーツ観戦と SNS のリアルタイム性は親和性が高く, 試合を見ながらたくさんの人が同時に投稿する傾向がある. ツイートには観戦者たちの試合に対する率直な感想が表れており, それらは出場選手のパフォーマンスの良し悪しを示唆すると考えられる. サッカーなどのスポーツの試合を振り返る指標として出場選手の採点というものがあり, 出場選手がその試合でどれほど活躍したかがわかる. その出場選手の採点において重要なのは客観性である. 採点に客観性を持たせるために観戦者のツイート内容を利用する. 本研究では特定の試合に関するツイートを集め出場選手ごとにツイートを分類し, それらを感情分析にかけることによって出場選手の採点を行う.

キーワード サッカー, スポーツ, 選手評価, Twitter, 感情分析

1 はじめに

近年, Twitter¹などの SNS が普及したことにより, その時々起きたことや感じたことをほぼリアルタイムに SNS に投稿する人が多く存在している. Tech Crunch²によると, Twitter Japan は 2018 年 12 月 26 日に事業戦略説明会を開催した [1]. その場で Twitter は全世界での月間アクティブユーザー数の発表があり, その数は 2018 年 10 月現在で 3 億 2600 万人を超えている.

さらに事業戦略説明会ではアンケートサイト Macromill³が行ったアンケート結果が発表された. このアンケートは, 2018 年 5 月に 15 歳から 64 歳のスマートフォン利用者を対象に行われていた. Twitter と名前は具体的には明記されていなかったが別の SNS 二つとメッセージングサービスを比較して, それぞれのサービスがどのような特色を持ったサービスであるかというアンケートになっていた. その結果によると Twitter は他 SNS と比べて災害, 交通機関の遅延, 旬なトピックの情報がすぐに手に入ると言われており, もっともリアルタイム性の高い SNS であると言える.

そして Sprinklr⁴による「2018 FIFA ワールドカップ ロシア」の開催期間中の調査では, SNS で交わされるワールドカップ関連の会話の約 86.3 % が Twitter で起きていたという結果が出ていることも発表された. このように, サッカーなどのスポーツ観戦と Twitter のリアルタイム性は親和性が高く, 試合を見ながらたくさんの人が同時に投稿する傾向がある.

また, サッカーなどスポーツの試合を振り返る為の指標とし

て選手の採点というものがある. 試合を見ていない人がどの選手が活躍したか簡単に知るための指標としても使われることがあるので, 採点に重要なのは客観性である. 野球のようにプレーが断続的なスポーツは一つ一つのプレーのデータから選手の活躍がわかりやすいが, サッカーのようなプレーが連続的なスポーツはデータには現れない活躍があり, データだけでは採点がしづらい.

また選手の採点はスポーツサイトやニュース記事でよく行われており, サイトの運営者やライターが試合を見て採点したものが載っているのが主だ. これらは, プロの意見とは言っても個人の意見であることが多いため, サイトや記事によって採点がばらけることがある.

しかし, Twitter には観戦者の試合に対する意見がたくさん投稿されており, 出場選手のパフォーマンスを示唆していると考えられる. 本研究では, それらのツイートを利用することで多くの観戦者の意見を反映した客観性のある出場選手の採点ができると考え, 採点システムの提案を行う. また, 本研究ではサッカーの試合での選手の採点を想定している.

サッカーの採点は国によって方式に違いがある. 主流な採点方式としては 10 点を満点として, 6 点を及第点として可もなく不可もなく平均的な活躍した選手に与える方式がある. 日本でもこの方式が多く使われている. 本研究でもこの方式を参考にしたシステムの提案をしている. 他にも 1 点から 6 点で採点し, 数字の小さい方が評価が高いという方式もある.

本研究では, ツイート内容から選手に対する意見や評価をスコア化するに当たって感情分析という技術を用いている. これは, テキストからユーザーの意見や評価を分析することができる. 感情分析は, 出来事や製品に関する意見を SNS などインターネットから効率的に収集することができる.

本研究では, 特定の試合に関するツイートを集め, そしてその試合に出場している選手ごとにツイートを分類し, それらに

1: <https://twitter.com/>

2: <https://jp.techcrunch.com/>

3: <https://www.macromill.com/>

4: <https://www.sprinklr.com/>

スコアを与えるために感情分析にかけることによって出場選手の採点をする。また本論文の構成は以下の通りである。2 章では、関連研究について述べる。3 章では、提案手法について述べる。4 章では、実験の結果と考察について述べる。5 章では、本研究のまとめと今後の課題について述べる。

2 関連研究

本研究のように、Twitter へのスポーツの試合やテレビ番組などの感想の書き込みを利用した研究は他にも行われている。小林ら [2] は Twitter に投稿された実況書き込みから視聴者が応援するチームに関する属性を判別し、同一チームを応援する視聴者の実況書き込みの状況から、視聴者の意見を反映した要約映像を生成した。

白鳥ら [3] は、今まで Twitter におけるサッカーのネタバレ情報のデータセットを構築し、特徴的な単語やその時間的特徴について明らかにしてきた。さらにスポーツの種類を拡張しデータセットを構築し直すことで、試合展開とネタバレの数や内容が大きく関係していることを明らかにした。

中澤ら [4] は対象番組の放送中に投稿された Tweet を収集し、一定時間単位の Tweet 数をカウントし、Tweet 数が急増した時間帯を重要シーンとして検出し、そのシーンにおける主要人物を推定し、その人物をもとにイベント内容を推定する手法を提案した。プロ野球番組を対象に提案手法を評価した結果、試合中の重要シーンを高精度で検出できることが示されている。

本研究では、ツイートにスコアを与えるために自然言語処理・テキストマイニング向けの Python ライブラリ NLTK(Natural Language Tool Kit)⁵の機能の Sentiment Analysis を用いたが、この機能では VADER という手法が感情分析のルールとして用いられている。この VADER は Gilbert ら [5] によって発表されている。VADER にはソーシャルメディアに強かったり、学習データがいらないなどの特徴がある。

3 提案手法

まず、図 1 に本研究の提案手法の流れを示す。はじめに、特定の試合に関するツイートであることを示すタグで指定しツイートを収集する。次に、収集したツイートをツイート内に含まれる出場選手の名前を用いて、各出場選手ごとに分類する。分類したツイートをそれぞれ感情分析にかけることによって各ツイートにスコアを与える。それらのスコアを用いて計算することによって各出場選手の採点を行う。

3.1 ツイートの収集

本研究では、収集方法として Twitter Streaming API を用いて、特定のサッカーの試合に関するツイートを収集する。特定の試合に関するツイートであることを示すタグを指定して検索することによって目的のツイートを収集する。

また本研究では、なるべくツイート数の多い試合で実験する

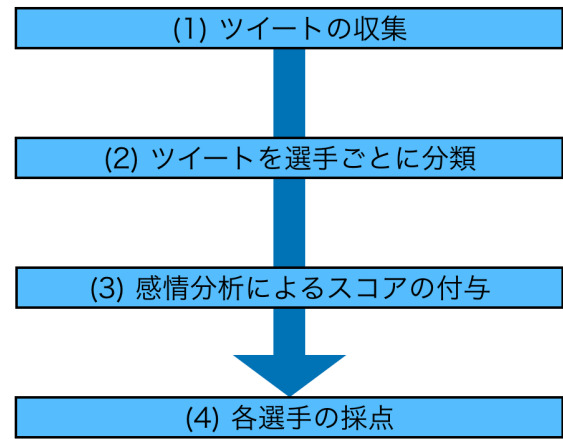


図 1 提案手法の流れ

ために、日本語ではなく英語のツイートを収集した。

3.2 ツイートを選手ごとに分類

本研究では、サッカーの試合を対象としているため、まず収集したツイートを試合開始時間からサッカーの一般的な試合時間の 2 時間で限定する。

今回の収集方法では、リツイートも一つのツイートとして集められてしまうので、全く同じ内容のツイートをリツイートされた分、重複して収集してしまう。リツイートとして収集されたツイートはテキストの先頭に 'RT' の文字とリツイート元のユーザー名が付与される。先頭に付与される 'RT' の文字を含むツイートを省くことによって、全く同じ内容のツイートが分析の対象になることを防ぐ。

このように、限定されたツイートの中から、各出場選手ごとにツイート内容に選手名を含んでいるものを抽出することによって、出場選手ごとにツイートの分類を行う。

3.3 感情分析によるスコアの付与

本研究では、NLTK の機能の Sentiment Analysis を用いる。例として 'I am happy' という文章に、この機能で感情分析を行うと、

```
'I am happy'
{'neg': 0.0, 'neu': 0.213, 'pos': 0.787, 'compound': 0.5719}
```

というような出力がなされる。このように否定的 ('neg'), 中立 ('neu'), 肯定的 ('pos') の値と VADER によって定義された総合的な感情値である 'compound' という値でテキストが評価される。'neg', 'neu', 'pos' は 0~1 で表され、'compound' は -1~1 で表される。本研究では、これらの値を用いて分類されたツイートに 3 通りのスコアを与えて各スコアで採点結果を比較した。3 通りのスコアは以下の通りであり、各スコアを式で表す。

(1) 各選手ごとにツイートの 'pos' と 'neg' の差を全て足した

⁵ : <http://www.nltk.org/>

もの

$$s_1 = \sum_{i=0}^{T_k} (v_{i_{pos}} - v_{i_{neg}}) \quad (1)$$

(2) 各選手ごとにツイートの'compound'の値の和を取ったもの

$$s_2 = \sum_{i=0}^{T_k} v_{i_{compound}} \quad (2)$$

(3) 各選手ごとにツイートの'compound'の値の和を各選手ごとにツイート数で割り平均を取ったもの

$$s_3 = \frac{\sum_{i=0}^{T_k} v_{i_{compound}}}{T_k} \quad (3)$$

ただし、 s_1 , s_2 , s_3 は (1), (2), (3) の方法で各選手に与えられるスコア、 T_k は各選手ごとの総ツイート数、 $(v_{i_{pos}}, v_{i_{neg}}, v_{i_{compound}})$ は各ツイートの'pos', 'neg', 'compound'の値とする。

3.4 各選手の採点

サッカーの選手の採点は 6.0 を平均にすることが多いため本研究でも計算の際に 6.0 を平均に扱い採点をする。各選手の採点を以下の式にて求める。

$$R_k = \frac{10(s_k - \mu)}{\sigma} + 6.0 \quad (4)$$

ただし、 R_k は選手 k の採点、 s_k は選手 k の 3 通りの方法で求めたスコアのうちの一つ、 μ は全選手のスコアの算術平均、 σ は全選手のスコアの標準偏差とする。以上のように、6.0 を平均とした偏差値のように求める。

本章で定義した記号とその意味を表 1 にまとめる。

3 通りのスコアから求めた採点を実際のスポーツサイトやニュース記事の採点と比較することで分析をする。

4 実験

4.1 データセット

本研究では、イングランドの 1 部リーグであるプレミアリーグから 18-19 シーズンの第 14 節アーセナル対トッテナム・ホットスパー (以下トッテナムという) と第 17 節リヴァプール対マンチェスター・ユナイテッド (以下マンチェスター U という) の 2 試合に関する英語のツイートをを用いた。ツイートの検索の際に用いたタグとしては、各チーム名を 3 文字に省略し、対戦す

表 1 第 3 章にて定義した記号と意味

記号	意味
s_1	(1) の方法で与えられる各選手のスコア
s_2	(2) の方法で与えられる各選手のスコア
s_3	(3) の方法で与えられる各選手のスコア
T_k	選手 k の総ツイート数
$(v_{i_{pos}})$	各ツイートの'pos'の値
$v_{i_{neg}}$	各ツイートの'neg'の値
$v_{i_{compound}}$	各ツイートの'compound'の値
R_k	選手 k の採点結果
s_k	3.3 で求めた選手 k のスコア
μ	全選手のスコアの算術平均
σ	全選手のスコアの標準偏差

表 2 各試合で使ったタグ

対戦カード	タグ
アーセナル 対 トッテナム	#ARSTOT
リヴァプール 対 マンチェスター U	#LIVMUN

る 2 チームでくっつけたものである。表 2 に本研究で用いたタグを示す。

これらの試合に関するツイート数は分析の対象となる試合時間の 2 時間でリツイートを除いたものはそれぞれ 23467 件と 21442 件となっている。これらの試合はプレミアリーグの中でも人気のチーム同士の対戦であるため他の対戦カードと比べても非常に多くなっている。表 3 にプレミアリーグの比較的人気の高いチームの試合を中心にいくつかの試合のツイートを集め、試合時間やリツイートを除いたツイート数を示す。

表 3 他試合とのツイート数の比較

対戦カード	ツイート数
アーセナル 対 トッテナム	23467
リヴァプール 対 マンチェスター U	21442
アーセナル 対 リヴァプール	14431
ウエストハム・ユナイテッド 対 マンチェスター U	12619
フルハム 対 アーセナル	6839
マンチェスター・シティ 対 サウサンプトン	2564

本研究で用いた試合は、比較的多量のツイートを集められる試合であると考えられる。

4.2 実験結果

提案手法により求められた各出場選手の採点とスポーツサイト WhoScored.com とニュースサイト skysports の記事による採点をグラフに表したものを図 2 と図 3 に示す。

また今後、アーセナル対トッテナム・ホットスパーの試合を試合 A、リヴァプール対マンチェスター・ユナイテッドを試合 B と表記する。

図 2 と図 3 において、skysports の採点で 0 点という評価の選手が両試合合わせて 7 人いるが、これらの選手は途中からの出場であり出場時間が短く試合にあまり関わることができなかったため、評価がなされていない。

3 通りのスコアによる採点と 2 つのサイトの採点とのそれぞ

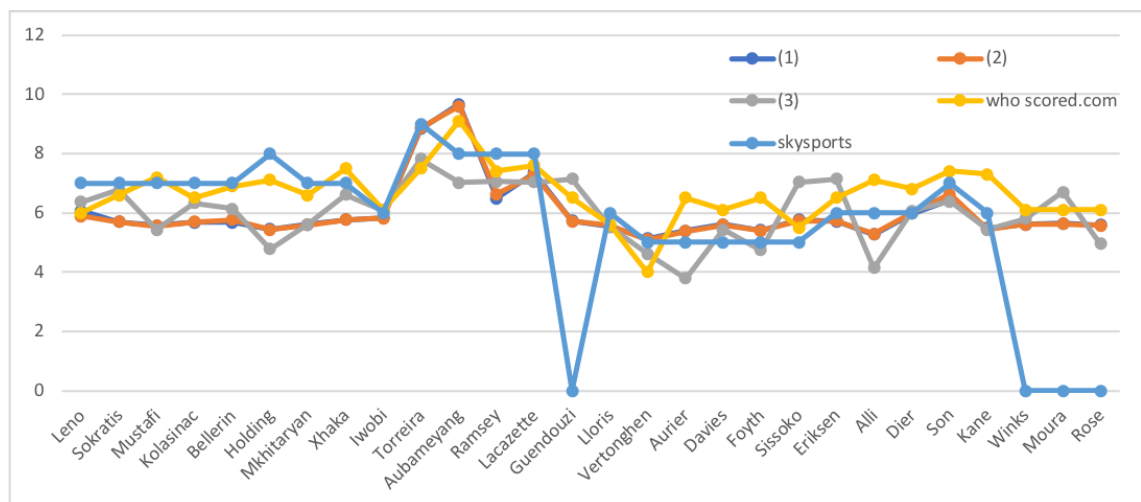


図 2 試合 A の各選手の評価

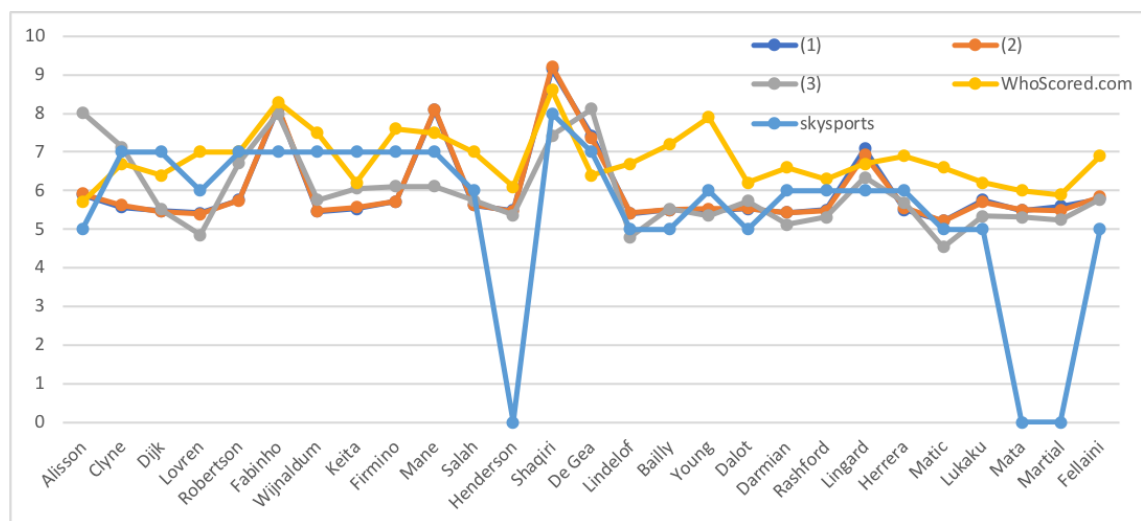


図 3 試合 B の各選手の評価

れの相関係数を計算し、表 4～9 に示す。また skysports の記事内において採点がなされていない選手がいるため、skysports で採点が行われている選手だけを用いて相関係数を計算した。さらに両チーム合わせた相関だけでなく勝利チームと負けチームで分けた時の採点の相関係数も計算をした。

また、二つのサイトの採点の相関係数を表 9 に示す。

表 4 (1) の方法と WhoScored.com の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.64184001	0.75854038	0.4654174
試合 B	0.55051353	0.73015803	-0.1881831

表 5 (1) の方法と skysports の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.65657713	0.68901147	0.64185631
試合 B	0.53889942	0.48074221	0.63372648

表 6 (2) の方法と WhoScored.com の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.65294013	0.76824846	0.4751149
試合 B	0.56634585	0.72864448	-0.1677127

表 7 (2) の方法と skysports の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.66558581	0.70135286	0.68788223
試合 B	0.54576082	0.48405588	0.64481304

表 8 (3) の方法と WhoScored.com の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.34429992	0.38321295	0.06449515
試合 B	0.1677427	0.16075368	-0.1485389

4.3 考 察

本研究の目的としては、客観的な選手の採点システムの作成であり、採点サイトの採点結果に近づくものではない。つまり相関係数を 1 に近づくことではない。しかし提案手法による採点結果が採点サイトとの相関が低いと提案手法による採点

表 9 (3) の方法と skysports の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.57255769	0.47238991	0.36305905
試合 B	0.48222253	0.01808371	0.68476451

表 10 二つのサイトの採点の相関係数

対戦カード	両チーム	勝利チーム	負けチーム
試合 A	0.671198	0.63042596	0.60302269
試合 B	0.51669764	0.65937919	0.05191279

は的外れな採点であることになってしまう。

どの程度相関があれば、選手の採点システムとして有用か調べるために基準として、WhoScored.com と skysports の採点の相関係数を計算した。表 10 に示されている相関係数に近い相関係数であれば、採点システムとしてかなり有用であると考えられる。

まず表 4～9 より、(3) の方法は (1), (2) の手法と比べて、相関係数が低くなっていることがわかる。(3) の方法のみスコアの計算時にツイート数で割り、平均を求めている。したがってツイート内容だけでなく、選手ごとのツイート数も選手を評価する指標になっていると考えられる。

さらに、両チームでの相関係数を比べると、試合 A では基準となる 2 サイト間の相関係数を超えるものはなかったが、(2) の方法が 2 サイト共に最も高い相関を持っていることがわかる。試合 B においては (1), (2) の二つの方法が基準の相関係数を上回り、こちらの試合も (2) の方法が最も高い相関を持っている。

次に、勝利チームと負けチームの相関を比べると、ほとんどの結果において勝利チームの方が相関が高くなっている。これは、勝利チームの方が負けチームよりもツイート数が多いことによる結果だと考えられる。

また、試合 A と試合 B で相関係数を比べると、試合 A の方が相関係数が高い傾向にある。表 2 を見ると、試合 A の方がツイート数が多い。ツイート数が増えるほど、採点システムの精度が上がると思われる。

5 まとめと今後の課題

本研究では、試合に関するツイートを集めツイート内容を感じ分析にかけることによる出場選手の採点システムを提案した。また、実際のスポーツサイトなどの採点との相関を求め、提案手法の有用性を分析した。その結果、本研究で提案した 3 手法の中だと (2) の各選手ごとにツイートの 'compound' の値の和で計算した手法が採点サイトとの相関がもっとも採点サイト同士の相関に近くなっており、もっとも有用なシステムであると考えられる。また、ツイート数が採点を左右するのでツイート数が多くなるような注目試合出ないと本研究で提案した手法は効果を発揮できない可能性が高いと思われる。

今後の課題として、今回基準とした表 10 の相関係数に近い相関があり的外れではない採点システムの手法を提案できたが、完全に超えられる手法を見つけれなかったため、より精

度が上がる手法を提案する必要があると考える。そのためには、NLTK の Sentiment Analysis に頼るのではなく、自ら単語のスコアリングを行い、感情分析をするのが一つの手であると考えた。特に、サッカーやスポーツに関連する単語に着目したスコアリングが重要であると考えた。スコアリングの際にプレーの種類に関する単語らの重要度を可変させて採点を行うことで各選手がどのようなプレーが優れているかを知ることができる。また、ツイート数が少ない試合でも高い精度で採点ができるシステムの提案も必要がある。

さらに各選手の採点の計算の際に、第 2 章の関連研究で取りあげたようなユーザー属性の判別 [2] を行うことによって、応援しているチームの選手へのツイートであれば若干の重みを持たせるなどの工夫や、時間帯や試合展開も [3][4] 考慮したスコアリングなどの工夫もできると考える。

また、採点結果の根拠も示すことが必要だ。各選手のどのプレーがいい評価、悪い評価に繋がったのかしっかりと表したい。例えば、選手ごとに選手名が含まれるツイートの中にもっとも多く現れた単語やフレーズの表示や、選手ごとにもっとも多くツイートが投稿された時間帯を表示することで各選手の採点結果の要因となったプレーや時間帯を示すことができるのではないかと考えられる。

今後の展望として、サッカーの試合の採点システムだけでなく、他のスポーツへの応用を検討している。また、今回は英語のツイートを対象に行なったが、ツイート数の少ない試合でも高い精度が出すことができれば日本語のツイートへの対応ができると考えられる。

文 献

- [1] Twitter の月間アクティブユーザー数は日本で 4500 万超。 <https://jp.techcrunch.com/2018/12/26/twitter-2/>. accessed: 2018-12-26.
- [2] 小林尊志, 野田雅文, 出口大輔, 高橋友和, 井手一郎, 村瀬洋. Twitter の実況書き込みを利用したスポーツ映像の要約. 電子情報通信学会技術研究報告, Vol. 100, No. 457, pp. 165–169, 2011.
- [3] 白鳥裕士, 中村聡史. スポーツジャンルに依るネットバレー特性分析と判定手法の提案. *DEIM Forum*, pp. B6–1, 2016.
- [4] 中澤昌美, 帆足啓一郎, 小野智弘. Twitter によるテレビ. 番組重要シーン検出及びラベル付与手法. *DEIM Forum*, pp. 517–519, 2011.
- [5] Gilbert and CJ Hutto Eric. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International Conference on Weblogs and Social Media (ICWSM-14)*. Available at (20/04/16) <http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf>, 2014.