

自己運動との比較による他者の指差し位置推定の 困難さの要因の検討

An Investigation on the Factors Contributing to the Difficulty of Interpreting Others' Pointing Gestures in Comparison with Self-Motion

峯 大典^{1*}, 中島 亮一² 浅野 倫子¹ 鳴海 拓志¹
Daisuke Mine, Ryoichi Nakashima, Michiko Asano, Takuji Narumi

¹東京大学, ²京都大学

¹The University of Tokyo, ²Kyoto University

* mine@cyber.t.u-tokyo.ac.jp

概要

本研究では、他者の指差し行動の解釈の困難さについて、視点の違いという従来の研究で示されてきた要因に加えて、運動の主体性という観点から検討した。一人称視点と三人称視点において、自己運動と他者運動の指差し行動から指差し位置を推定する課題を行ったところ、どちらの視点でも自己運動時の推定がより安定しており、視点によって正確性が異なった。つまり、指差しコミュニケーションの困難さが複合的な要因から生じていることが示唆された。

キーワード：指差し位置推定、自己・他者運動、一人称・三人称視点

1. 序論

指差し行動は一般的に用いられるコミュニケーションの一種であり、特に他者に対して口頭のみで伝えるのが困難な空間上の位置を伝える際によく用いられる。「あの赤い車を見て」というように言語を指差しと併用することもあるが、それが難しい場合（赤い車が並んでいる等）もある。しかし、指差しのみによって特定の位置を正確に伝えることは思いのほか困難であり、指差しの行為者の意図とは異なる位置に観察者の注意が向けられてしまうコミュニケーションエラーがしばしば生じる。近年、指差し行動におけるコミュニケーションエラーに関して、行為者と観察者の視点の違いに基づく説明が試みられている[1-3]。例えば Herbort et al.[3]では、実験参加者が、遠方の特定の位置（目標位置）を自己の指で示す指差し課題と、他者の指差し行動を観察し指差している位置を回答する観察課題を、それぞれ一人称視点、三人称視点で行った。その結果、一人称視点の場合、目標位置が目と指を結ぶ直線の延長線（eye-finger line: 以降 EFL）上にくるように指差しを行い、観察課題においてもその直線上に目標位置があると回答した。一方、三人称視点

の場合は、目標位置が肩と指を結ぶ直線の延長線

（shoulder-finger line: 以降 SFL）上にくるように指差しを行い、観察課題においても同様の回答をした。つまり、指差し行動やその解釈は、視点によって変化し、それが行為者と観察者間のコミュニケーションエラーの要因だと考えられる。

しかしながら、視点以外にも、指差しの行為者と観察者間のコミュニケーションエラーを生み出す要因はありうる。例えば、行為者は指差しを行う前に、あらかじめどこを指し示すかというトップダウンの意図を持ち、それをもとに自身の身体をコントロールして指差し行動を行う。一方、観察者はこの意図を事前には知ることができず、行為者の身体の動きを観察するだけである。このような運動の主体性の有無が、観察者にとっての指差し位置の推定を難しくしている可能性が考えられる。これを検証するために、本研究では、他者の指差し行動の観察に基づく空間位置推定の正確性および安定性について、自己の指差し行動の観察に基づく推定と比較する。本研究では、実際の指差し位置を、一人称視点では EFL 上、三人称視点では SFL 上にあると見なし、それと推定された指差し位置の差分（これを推定誤差と呼ぶ）の平均値を正確性の指標（0 に近いほど正確）とし、分散を安定性の指標（小さいほど安定）とする。

実験では、実験参加者が VR を用いて、一人称視点および三人称視点における自己と他者の指差し運動を観察し、指差し位置の推定を行った。指差し行動の解釈が EFL や SFL という物理的に一意に定まるルールのみ依存するのであれば、自己運動と他者運動による指差し位置の推定の正確性、安定性に差はないと予想される。一方、運動の主体であることが影響するのであれば、自己運動と他者運動で正確性、安定性の少なくとも一方には差が見られると予想される。また、指差し位置の推定が指差し対象との距離に応じて変わ

るのかも検討した。

2. 実験 1

VR 空間内のアバターの指差し運動を観察し、前方の壁面上のどこを指しているかを推定する課題を行った。アバターを自分が動かす条件（自己条件）、他者が動かす条件（他者条件）を設け、推定の正確性、安定性を比較した。実験 1a では一人称視点で、実験 1b では三人称視点でアバターを観察した。

参加者

実験 1a,1b に 16 名ずつ（実験 1a: 25.3±5.8 歳、女性 7 名; 実験 1b: 25.7±8.4 歳、女性 10 名）が参加した。全員が正常な視力または矯正視力を有しており、運動能力に支障がなかった。

実験装置・刺激・手続き

VR 提示装置として HTC VIVE（両眼解像度 2,160 × 1,200 px）を使用し、VR 環境は Unity 3D で開発した。

実験 1a では、参加者はアバターの運動を一人称視点で観察し、VR 内でアバターが指差ししている前方の壁面上の位置をマウス操作で回答した。自己条件では、参加者は右手人差し指で壁面上の任意の場所を指差し、それと連動したアバターの指差し運動を観察した。他者条件では、参加者とは別の実験協力者が右手人差し指で壁面上の任意の場所を複数回指差した運動を記録し、それをアバターの運動に反映させた。アバターと壁面との距離は 4 m、12 m の 2 条件あり、各参加者は他者・4 m、他者・12 m、自己・4 m、自己・12 m の 4 条件をそれぞれ 20 試行ずつ行った。各条件において、参加者の利き目と指を結んだ EFL と壁面の交点を各試行における原点とし、その原点と推定位置の差分（推定誤差）の平均値、分散を計算した。

実験 1b では、基本的な実験手続きは実験 1a と同じだったが、参加者はアバターの運動を三人称視点で観察した。アバターは壁に正対しており、参加者はアバターの後方 50 cm、左 50 cm 離れた視点からアバターを観察した。また、各試行における指差し位置の原点を、EFL ではなく SFL に基づいて算出した。

結果と考察

各実験の結果（水平方向）を図 1, 2 に示す。以降の検定は全て Wilcoxon の符号順位検定を用いた。実験 1a では、水平方向に関しては壁面の距離 4 m、12 m 条件で共に他者条件よりも自己条件の推定誤差の平均値が小さい、つまり EFL に近い位置を推定していた（4

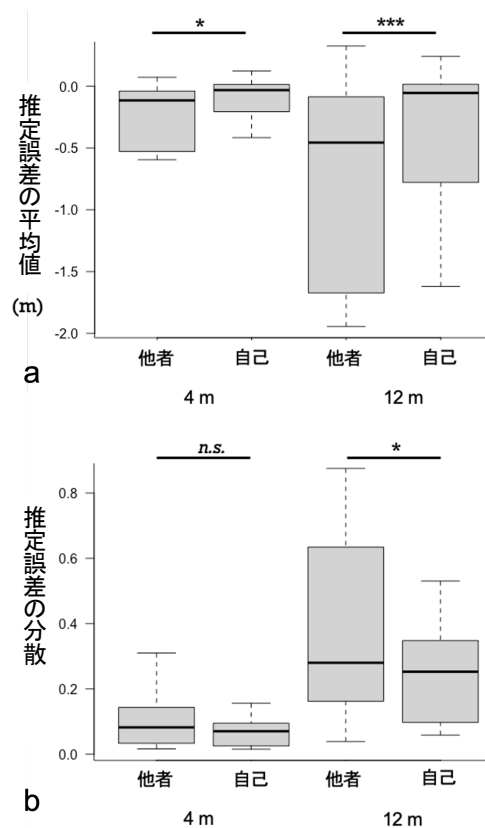


図 1: 実験 1a（一人称視点）水平方向における指差し位置の推定誤差の平均値 (a)、および分散 (b)。a の縦軸の正の値は基準点から左方向のずれを示す。

m: $V=6, p<.001$; 12 m: $V=20, p=.011$)。また、壁面 4 m 条件では自己条件と他者条件間で推定誤差の分散に有意差は見られず、壁面 12 m 条件では自己条件の方が他者条件より分散が有意に小さかった (4 m: $V=98, p=.130$; 12 m: $V=115, p=.013$)。一方、垂直方向に関しては自己条件と他者条件で平均値、分散ともに差が見られなかった (平均値 4 m: $V=74, p=.782$; 12 m: $V=80, p=.562$; 分散 4 m: $V=57, p=.597$; 12 m: $V=78, p=.632$)。

実験 1b では、水平方向に関して、壁面条件によらず、自己条件よりも他者条件の推定誤差の平均値が小さい、つまり SFL から近い位置を推定していた (4 m: $V=113, p<.001$; 12 m: $V=113, p<.001$)。実験 1a と同様、壁面 4 m 条件では自己条件と他者条件間で水平方向の推定誤差の分散に有意差は見られず、壁面 12 m 条件では自己条件の方が他者条件より分散が有意に小さかった (4 m: $V=83, p=.464$; 12 m: $V=111, p=.025$)。一方、垂直方向に関しては自己条件と他者条件間で平均値、分散ともに違いが見られなかった (平均値 4 m: $V=106, p=.051$; 12 m: $V=99, p=.117$; 分

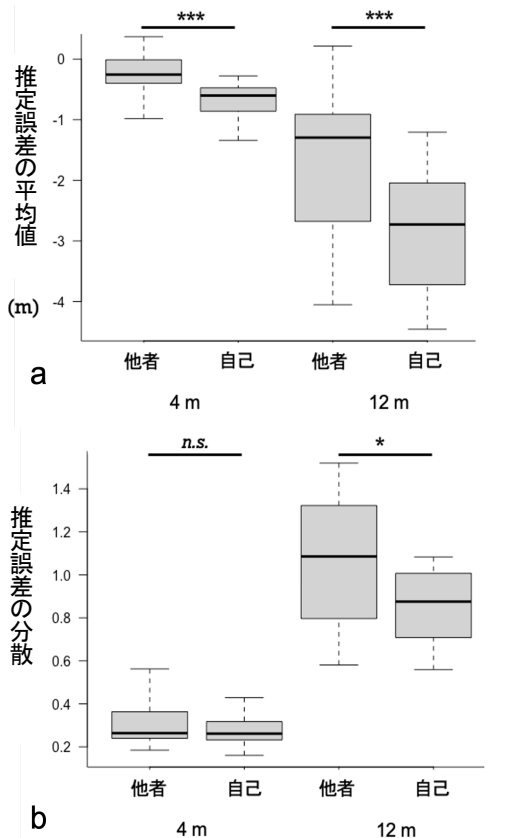


図2: 実験1b (三人称視点) 水平方向における指差し位置の推定誤差の平均値 (a)および分散 (b)。aの縦軸の正の値は基準点から左方向のずれを示す。

散 4 m: $V = 30, p = .051$; 12 m: $V = 49, p = .348$ 。

一人称視点の場合は、自己の指差し運動から推定された位置は、他者の指差し行動からの推定位置よりもEFLで物理的に定まる位置に近い、つまり正確であった。一方、三人称視点では、他者の指差し運動から推定された位置が、自己運動から推定された位置よりもSFLで定まる位置に近かった。具体期には、自己条件の推定位置が他者条件よりも右側により大きくずれていた。これは、視点に依存して運動の主体による指差し位置推定の正確さへの影響が異なることを示唆する。Herbort et al. [3]は、第三者視点での指差し位置の推定において、SFLに基づいた推定に加えて、自身の目線と他者の指を結んだ直線 (三人称EFL) に基づいた推定が影響をしていると報告している。実験参加者がアバターを左後方から観察していたことを考えると、自己運動で操作されたアバターを第三者視点で観察すると三人称EFLによる指差し位置推定の影響が大きくなったと解釈できる。

また、実験1aと1bに共通して、指差し対象が遠い場合に、他者の指差し位置の推定誤差の分散が大きく

なる、つまり安定性が低下する。よって、観察した指差し運動から推定される指差し位置はある一定の範囲をもっており、それは距離に応じて拡大していくと考えられる。ただし、自己が運動する場合と他者が運動する場合で、その範囲の拡大具合が異なり (他者条件でより大きく拡大する)、それが指差しコミュニケーションにおけるエラーの一因だと考えられる。

3. 実験2

実験1では、壁面が遠方にある場合に、指差し位置推定の安定性が他者条件の方が低かった。これは運動の主体性の有無の影響だと考えられるが、それだけではなく記憶の観点からも説明可能である。つまり、自己条件では、実験参加者は指差し行動を行う前に目標点を決定しており、その目標地点を正確に記憶できていれば、アバターの運動を観察しなくても指差し位置を推定可能である。実験2では、自己条件における目標位置の記憶が指差し位置推定の安定性に与える影響を調べた。そのために、一人称視点、三人称視点のそれぞれについて可視条件と不可視条件を用意した。可視条件は実験1の自己条件と同じであった。不可視条件では実験参加者が指差し行動を行った後、指差し位置の推定を行う前にVR内のアバターが消失した。この条件では、アバターを観察しながら指差し位置を推定することができず、事前の目標位置の記憶に強く依存して回答を行うことになる。もし指差し目標位置の記憶に依存した位置推定をしているならば、可視条件と不可視条件間の推定に差がないと予想される。

参加者、実験装置、刺激および手続き

実験参加者は16名の大学生、大学院生 (26.3歳 ± 7.9歳、女性10名)であった。全員が正常な視力または矯正視力を有しており、運動能力に支障がなかった。装置と刺激は実験1と同じであった。

各実験参加者が可視条件、不可視条件の両方を行った。各条件は、一人称視点と三人称視点のブロックに分かれており、条件の順序および各条件内でのブロックの順序はカウンターバランスした。アバターは常に参加者の運動に同期して運動した。アバターと壁面の距離は12 mのみであった。可視条件では、自分が操作したアバターの指差しを観察しながら、指差し位置をマウスクリックで回答した。不可視条件では、まず実験参加者はある壁面上の位置を指差し、その後アバターが消失してから、指差し位置を回答した。

結果と考察

実験2の結果(水平方向)を図3に示す。一人称視点では水平方向、垂直方向共に可視条件の方が不可視条件より推定誤差が小さかった(水平: $V=114, p=.016$; 垂直: $V=3, p<.001$)。一方で三人称視点では、垂直方向の推定位置の差は見られなかったが($V=40, p=.159$)、水平方向に関しては不可視条件の方が可視条件よりSFLに近い位置を推定していた($V=20, p=.011$)。アバターの手が不可視の場合、三人称EFLの影響が減少し、正確性が向上したと考えられる。

また、一人称視点では、水平方向と垂直方向ともに可視条件の方が不可視条件より分散が小さい、つまり安定性が高かった。よって、記憶の効果が支配的ではないことが示唆された(水平: $V=21, p=.013$; 垂直: $V=0, p<.001$)。一方で、三人称視点ではいずれの方向に関しても可視条件と不可視条件との間で有意差はなく(水平: $V=35, p=.093$; 垂直: $V=91, p=.252$)、自身の指差し行動を観察する場合に記憶に基づく推定により安定性が向上した可能性は否定できなかった。つまり、一人称視点と三人称視点では、異なる方略で指差し位置推定を行っている可能性がある。ただし、

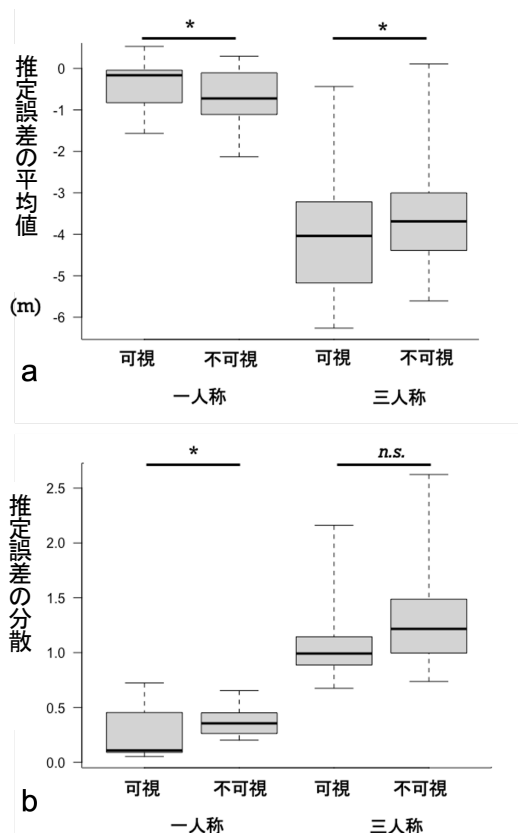


図3: 実験2の水平方向における指差し位置の推定誤差の平均値(a)、および分散(b)。aの縦軸の正の値は基準点から左方向のずれを示す。

少なくとも一人称視点では、運動の主体であることが指差し位置の推定に関与すると考えられる。

4. 総合考察

本研究では、VR装置を用いて、指差し運動の観察に基づく指差し位置の推定について、視点と対象との距離の影響を調べた。その結果、一人称視点と三人称視点のいずれの場合においても、自己運動の方が他者運動よりも位置推定の安定性が高かった。また、少なくとも一人称視点に関しては記憶以外の運動の主体であることの影響があることが示唆された。

運動の主体性については様々な解釈が考えられる。例えば、自分が操作する(すなわち運動主体感を伴う)身体は、他者が操作する身体に比べ空間位置の推定における手がかりとして働く可能性がある。また、運動しているのが自分自身であるという感覚(身体所有感)も、空間位置の認知に影響するかもしれない。運動の主体性のどのような要素によって指差し推定が変わるのか、今後詳細に検討する必要がある。また、一人称視点と三人称視点では正確性や安定性に関して部分的に異なる傾向があり、それぞれの視点では異なる方略を用いて推定がされている可能性もある。

指差しコミュニケーションでは、行為者の意図とは異なる位置に観察者の注意が向けられてしまうエラーが生じうる。先行研究では、視点の違いによるEFLとSFLに基づく指差し位置推定が原因であると報告されていた[1-3]。本研究は、それに加えて、運動の主体の違い(自己の指差し運動か他者の運動か)も指差し位置推定に影響することを示した。つまり、指差しコミュニケーションにおける行為者と観察者の齟齬は、複合的な要因によって生まれていると考えられる。

謝辞

本研究の一部はJSTムーンショット型研究開発事業(JPMJMS2013)の支援を受けて実施された。

文献

- [1] Wnuczko M, Kennedy JM. Pivots for Pointing: Visually-Monitored Pointing Has Higher Arm Elevations Than Pointing Blindfolded. *J Exp Psychol Hum Percept Perform.* 2011;37(5), pp. 1485-1491.
- [2] Herbert O, Kunde W. Spatial (Mis-)interpretation of pointing gestures to distal referents. *J Exp Psychol Hum Percept Perform.* 2016;42(1), pp. 78-89.
- [3] Herbert O, Krause L-M, Kunde W. Perspective determines the production and interpretation of pointing gestures. *Psychon Bull Rev.* 2021;28, pp. 241-248