Emotional Expressions Reconsidered

Challenges to Inferring Emotion From Human Facial Movements

Barrett, Adolphs, Marsella, Martinez & Pollak (2019)
Psychological Science in the Public Interest, 20(1), 1–68

Graduate Seminar in Affective Computing · 80 minutes

The claim, compressed

People do smile when happy and scowl when angry more often than chance — but not reliably enough, not specifically enough, and not consistently enough across cultures and contexts to read emotion off a face.

So a system that maps facial configurations onto emotion labels is not measuring emotion. It is measuring facial movement, and then guessing.

ROADMAP

How we'll spend the next 80 minutes

1	The common view and why it matters here	10 min
2	Reverse inference and the four criteria	15 min
3	Measurement: FACS, automation, ground truth	12 min
4	The production evidence	12 min
5	The perception evidence	13 min
6	Implications for affective computing	18 min

Three discussion breaks are built in. Slides are dense on purpose — skip freely.

MOTIVATION

Why this paper belongs in an affective computing seminar

Four of the five authors are not critics from outside the field.
Barrett is a constructionist emotion theorist. Adolphs is an affective neuroscientist. Marsella builds virtual humans. Martinez builds computer vision systems for facial action detection — he ran the EmotioNet Challenge. Pollak studies emotional development.

This is partly a critique of the field's own measurement stack, written by people who built parts of it.

- The inference these systems make is the paper's target
- The paper's negative result is a research agenda for us
- It names the accuracy cliff between lab and wild

Affectiva, Microsoft Emotion API, and peers

Marketed as detecting a fixed set of emotion categories, on the premise that these are universally signalled by particular facial configurations.

TSA / FBI screening

Agents trained to infer intent and deception from facial configurations, in programmes such as SPOT.

Clinical and legal use

Autism assessment and training tasks; jurors and judges reading remorse from a defendant's face.

Barrett et al. (2019), Introduction; author affiliations.

DEFINITION

The "common view" the paper sets out to test

Each of a small set of emotion categories is expressed by a distinctive configuration of facial muscle movements, reliably enough and specifically enough that observing the configuration diagnoses the emotional state — the way a fingerprint identifies a person.

What almost nobody defends

That every instance of anger is physically identical. Keltner and Cordaro concede there is no one-to-one correspondence between muscle actions and emotional experience.

The basic emotion position

Instances vary around a prototype. Deviation is attributed to display rules, regulation, cultural dialects, or noise — the prototype still diagnoses the state.

The constructionist position

Variation is by design, tailored to context: the situation, the body's internal state, the relationship, the culture. There is no prototype to recover.

Barrett et al. (2019), "The Common View"; Keltner & Cordaro (2017); see the paper's Figure 1 and Box 2.

STAKES

Where the common view is already doing work

Commercial emotion AI

Vendors sell facial emotion detection to advertisers, hiring platforms, and market researchers.

Border and aviation security

Screening programmes trained agents to read intent and deception from facial behaviour.

Courtrooms

Remorse and credibility inferred from a defendant's face; perceived untrustworthiness tracks harsher sentences.

Psychiatric assessment

Emotion-matching tasks used in autism research; training that generalises poorly to real interaction.

Early education

Feeling charts, books, and children's television teaching a canonical face per emotion.

Adjacent sciences

Neuroscience, computer vision, and AI inherit the six-category assumption in their stimulus sets.

Barrett et al. (2019), Introduction.

DISCUSSION 1

~5 minutes

Before we look at the evidence: what would it take to change your mind?

- If you have built or used an emotion recognition system: what did you take your labels to mean? A felt state, a display, or a facial configuration?
- What accuracy number would you need to see before you'd deploy such a system in hiring? In airport screening? Does the threshold change with the stakes?
- Whose ground truth were you validating against — self-report, actor intent, or another annotator's judgement of the same face?

1

Reverse inference and the four criteria

The conceptual apparatus that does the work in this paper

CORE CONCEPT

Two different questions, routinely conflated

FORWARD INFERENCE

p (facial configuration | emotion)

Given that someone is angry, how likely are they to scowl?

This is what production studies measure.

REVERSE INFERENCE

p (emotion | facial configuration)

Given that someone is scowling, how likely are they to be angry?

This is what every deployed system does.

Evidence for the forward direction does not license the reverse. Getting from one to the other requires the base rate of the emotion and the base rate of the configuration — and the configuration's false positive rate, which most published studies never report.

Barrett et al. (2019), "A Systematic Approach for Evaluating the Scientific Evidence"; see Figure 3.

FRAMEWORK

Four criteria a facial configuration must meet

Reliability

Evaluated

Does the configuration occur when the emotion occurs, more often than chance? Consistency of the association.

Specificity

Evaluated

Does it occur only for that emotion — and not for other emotions, or for non-emotional states? Uniqueness.

Generalizability

Evaluated

Does the pattern replicate across cultures, ages, methods, and populations — infants, blind individuals, remote societies?

Validity

Almost never addressed

Is the person actually in the emotional state? Requires a measure of emotion independent of the face itself.

Barrett et al. (2019), Table 2; criteria adapted from Haidt & Keltner (1999).

FRAMEWORK

The same criteria as a detection problem

	PERSON IS ANGRY	PERSON IS NOT ANGRY
SCOWL OBSERVED	**True positive** Reliability lives here. Widely reported.	**False positive** Specificity lives here. Almost never reported.
NO SCOWL OBSERVED	**False negative** Anger expressed some other way.	**True negative** Correct rejection.

The literature has thoroughly characterised one cell of this table and largely ignored the one that determines whether the inference works.

Barrett et al. (2019), Figure 3 and Table 2, notes 8–9.

STANDARDS

What counts as strong evidence

Unsupported	Weak	Moderate	Strong
below 20%	20 – 39%	40 – 69%	70 – 90%

Where production evidence lands

Spontaneous facial movement during emotional episodes sits in the weak band, and below it for infants and congenitally blind individuals.

Where perception evidence lands

Strong — but only in the specific task that supplies participants with a short menu of emotion words to choose from.

A correlation of r ≈ .20–.39 means a person sometimes scowls in anger — not most of the time. Weak reliability is itself evidence that unmeasured factors are driving the face. The authors call those factors context.

Barrett et al. (2019), Table 2; thresholds from Haidt & Keltner (1999), Ekman (1994), Russell (1994).

ALTERNATIVE HYPOTHESIS

What "context" actually stands for

Most studies test the common view against chance. Very few test it against a specific alternative — that facial movements are systematically shaped by measurable factors.

Situational

At work, at school, at home; what the situation affords

Social

Who else is present, and the relationship to them

Internal physical

Metabolic state, fatigue, hunger, illness

Internal mental

What is remembered, expected, appraised

Temporal

What happened a moment ago; where in the episode

Cultural

Individualist or collectivist, open or rule-bound

When context is studied at all, it is usually cast as a moderator of a universal expression — which preserves the assumption rather than testing it.

Barrett et al. (2019), "The Null Hypothesis and the Role of Context"; Boxes 2–3, SOM.

2

Measurement

FACS, automated coding, and the ground truth problem

INSTRUMENTATION

FACS: the measurement layer everything rests on

- Anatomically grounded: action units map to muscle contractions, not to emotion labels
- Expert human coders reach inter-rater reliability near .80 for individual AUs
- Descriptive by design — Rosenberg and others prefer "facial events" precisely because most facial behaviour expresses no internal state

The slippage happens one level up: AUs are anatomical, but the configurations assembled from them are then named as emotions. FACS itself makes no such claim.

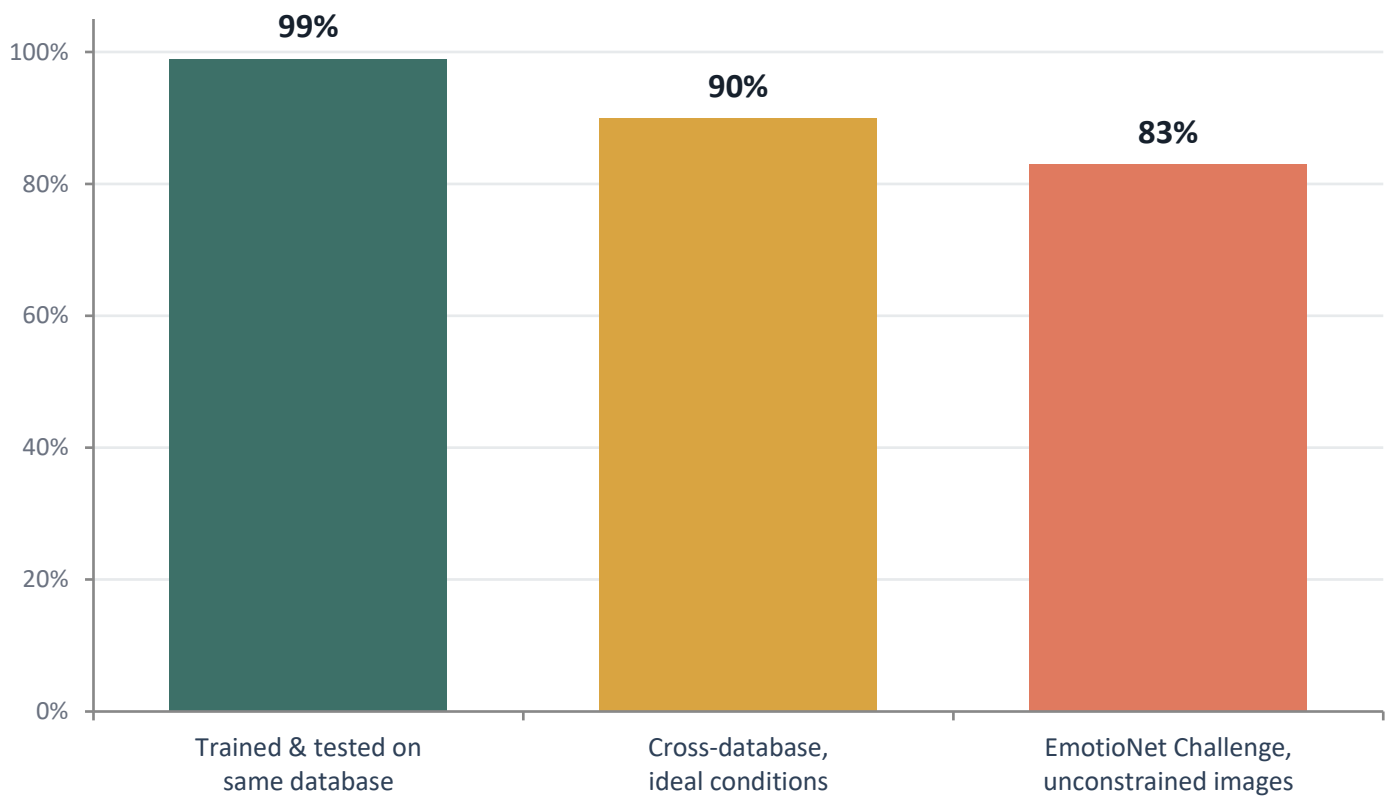
A FEW ACTION UNITS

AU 4	**Brow lowerer**	*Corrugator supercilii*
AU 6	**Cheek raiser**	*Orbicularis oculi (orbitalis)*
AU 9	**Nose wrinkle**	*Levator labii sup. alaquae nasi*
AU 12	**Lip corner puller**	*Zygomaticus major*
AU 15	**Lip corner depressor**	*Depressor anguli oris*

Barrett et al. (2019), Table 3 and "Measuring facial movements"; Ekman & Friesen (1978); Jeni, Cohn & De la Torre (2013).

AUTOMATION

Automated AU coding: the drop from lab to world



F1 below .65

In the challenge, the combined precision-and-recall measure fell below .65 on unconstrained images.

38 entered, 4 finished

Thirty-eight groups declared for the 2017 challenge, trained on one million images. Four submitted results.

The ceiling is soft

Manual verification of the reference images was itself estimated at around 81%.

Barrett et al. (2019), "Measuring facial movements with automated algorithms"; Benitez-Quiroz et al. (2016, 2017a); note 17.

VALIDITY

The ground truth problem

Suppose AU detection were perfect. You would still need to know what the person actually felt.

Autonomic physiology *Heart rate, skin conductance, respiration*	Meta-analyses find no consistent, category-specific autonomic signature. In anger, blood pressure can rise, fall, or hold steady — and a rise is not unique to anger.
Central nervous system *Neural activity, blood flow*	Individual studies report discriminating patterns; those patterns do not replicate across studies, even with matched methods and populations.
Self-report *Asking the person*	A verbal categorisation of experience, not a direct readout of state — and it uses the same folk categories under test.
Elicitation condition *Assuming the manipulation worked*	Assumes the film clip or scenario produced the intended state in every participant.

Barrett et al. (2019), "Measuring an emotional state"; Siegel et al. (2018); Box 8, SOM.

DISCUSSION 2

~6 minutes

If there is no independent measure of emotional state, what exactly is an emotion classifier learning?

- Is "emotion recognition" a supervised learning problem with a corrupted label set, or a task that has been misdescribed from the start?
- Would a perfect AU detector fix any of this? What would it fix, and what would it leave untouched?
- Is predicting annotator consensus a legitimate objective, if we say that is what we are doing? What could such a system honestly be used for?

3

The production evidence

How do people actually move their faces during emotional episodes?

PRODUCTION

Spontaneous facial movement in the laboratory

An expanded meta-analysis of studies observing how people moved their faces when exposed to emotion-eliciting events:

47

studies, 89 effect sizes, 3,599 participants

r = .32

average correlation between configuration and emotion (range .25–.38)

.19

average proportion of emotional episodes showing the predicted configuration (range .15–.25)

not reported

no overall specificity estimate: published studies omit the false positive rate

Every category except fear exceeded chance. All sat in the weak band. More often than not, people moved their faces in ways the common view does not predict.

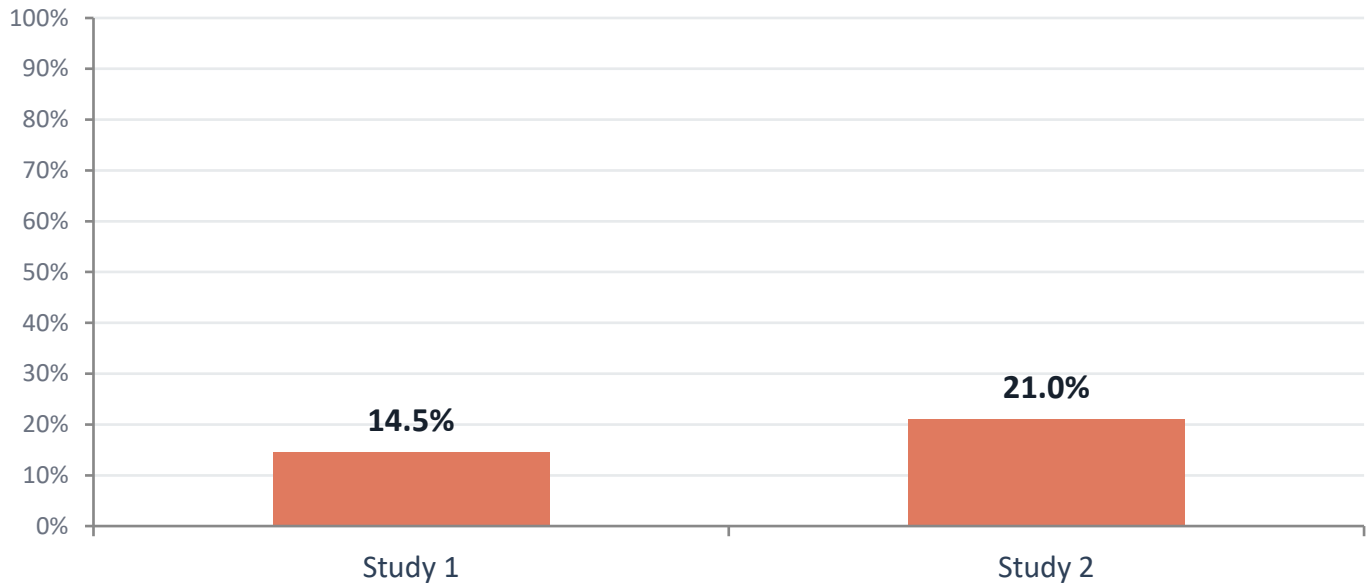
A specificity failure by name: the Duchenne smile, long treated as the marker of authentic happiness, can be produced deliberately when unhappy, and often signals submission or affiliation instead.

Barrett et al. (2019); Duran et al. (2017); Duran & Fernández-Dols (2018); Gunnery & Hall (2014); Rychlowska et al. (2017).

PRODUCTION

Out of the lab: winning judo fighters

Two field studies of Olympic and international judo competitors, photographed at the moment of victory — arguably the cleanest naturalistic happiness elicitation available.



Study 1

Eight of 55 winning fighters produced a Duchenne smile. Every one occurred during a social interaction.

Study 2

Twenty-five of 119 winning fighters produced one.

What predicted the smile

Whether an audience was being addressed — not self-reported happiness after the win.

Barrett et al. (2019); Crivelli, Carrera & Fernández-Dols (2015).

PRODUCTION

Posed configurations and remote cultures

Posing produces the strongest support — and means the least

When people are given an emotion word and asked to pose how they would express it, the canonical configurations appear, sometimes strongly. But this measures shared beliefs about expression, not how people spontaneously move. The authors read it as consensus, not evidence for a mechanism.

Small-scale, remote societies

The evidence for spontaneous production is unclear rather than negative — there is very little of it. The foundational cross-cultural work of the 1970s, including the Fore studies, did not systematically measure how people in these communities actually moved their faces during emotional episodes.

A structural problem worth noticing

The stimulus sets used across this literature — including the ones our datasets inherit — descend from a small number of posed photograph collections, selected decades ago for maximal distinguishability. They are caricatures, chosen because they are easy to categorise. Support for the common view rises with how posed and exaggerated the stimulus is.

Barrett et al. (2019), production sections and Table 5; Gendron & Barrett (2017); Russell (1994).

PRODUCTION

The two populations that should show it most clearly

INFANTS AND YOUNG CHILDREN

- Reliability and specificity are unsupported: the hypothesised configurations are not what infants reliably do
- The proposed fear configuration is rarely observed in young infants at all
- Infants scowl while crying and when about to cry — not distinctively in anger
- In one naturalistic study, children playing a deliberately frightening game produced the predicted fear configuration only rarely
- Adult observers still judge infants as angry or fearful — the perceiver supplies the category

CONGENITALLY BLIND INDIVIDUALS

- The strongest available test of an innate, non-imitative expressive programme
- Blind and sighted individuals moved their faces in broadly similar ways during emotional episodes
- But both groups produced the hypothesised configurations with weak reliability or none, and neither with specificity
- Similarity between the groups is often reported as support for universality — it is equally consistent with both groups simply being variable
- Onset and severity of blindness vary widely across these studies

Barrett et al. (2019), infant, child and congenital blindness sections; Camras et al.; Galati et al. (1997, 2001, 2003); Valente et al. (2017).

INTERIM SUMMARY

Where the production evidence leaves us

Adults across cultures, infants and children, and congenitally blind individuals all show substantially more variability in how they move their faces during emotional episodes than the common view allows.

Many-to-many, not one-to-one

Anger is expressed with a wider range of movements than a scowl, and scowls express more than anger.

A scowl, not the scowl

A scowl can be part of an instance of anger. It is not the expression of anger in any generalisable way.

Stereotype, not prototype

A prototype is a category's most typical member. A stereotype is an oversimplification treated as more general than it is.

Terminological consequence: prefer facial configuration, pattern of facial movements, or facial actions over emotional expression, which smuggles in the conclusion.

4

The perception evidence

Do people reliably infer emotion from facial movements?

PERCEPTION

The task you choose determines the answer you get

CHOICE FROM ARRAY

Participant sees a face and picks from a short list of emotion words supplied by the experimenter.

- Produces strong, replicable agreement
- Supplies the categories it then confirms
- Chance is high — six options, not open-ended

FREE LABELLING

Participant volunteers whatever word best captures the face — emotion word or not.

- Agreement falls to weak or moderate
- Specificity, where measured, is weak
- Participants often volunteer actions, not states

The most striking demonstration: the choice-from-array design produces apparent recognition even for invented emotion categories paired with made-up expressive cues. The task itself manufactures agreement.

Barrett et al. (2019), Table 5 and "Summary of Scientific Evidence on the Perception of Emotion in Faces".

PERCEPTION

The evidence base that built the consensus

The most recent meta-analysis of emotion perception studies remains the 2002 summary by Elfenbein and Ambady.

87
experiments summarised

22,000+
participants

20+
cultures sampled

4
studies using spontaneous rather than posed faces

The headline result

Within-culture agreement was strong; cross-culture agreement moderate — the ingroup advantage.

The design

Ninety-five percent of studies used posed configurations; all but four used choice from array.

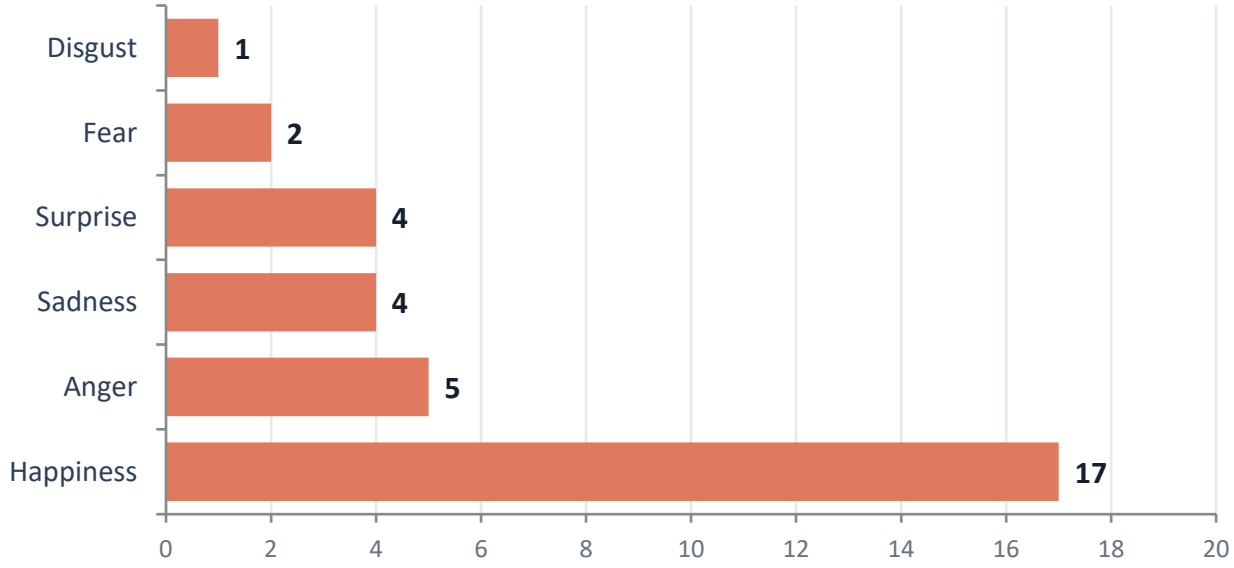
Most of these studies never report how often a scowl was perceived as something other than anger. Reliability without specificity does not license the reverse inference.

Barrett et al. (2019); Elfenbein & Ambady (2002).

PERCEPTION

Free labelling reveals a many-to-many mapping

Speakers of English, Spanish, Mandarin, Farsi, Arabic and Russian freely labelled 35 cross-culturally identified facial configurations. Reliability was moderate. Specificity was absent — multiple distinct configurations were labelled with the same emotion word.



Seventeen faces for happiness

If seventeen distinct configurations are all labelled happiness, no single one of them is the expression of happiness — and observing any one of them tells you little.

Not just cross-cultural noise

The variability appears within a single culture and a single language, not only across them.

Beyond dialects

This exceeds what accent or display-rule accounts predict.

Barrett et al. (2019), Figure 8b; Srinivasan & Martinez (2018).

PERCEPTION

Small-scale, remote societies

HADZA · TANZANIA

- Hunter-gatherers in an ecology resembling the one in which universal expressions are said to have evolved
- Above-chance matching occurred mainly when the target and foil differed in pleasantness — valence, not category
- Of 27 participants with minimal outside contact, only 12 matched the wide-eyed gasping face to the fear story above chance
- Participants with schooling or Swahili chose the canonical face more consistently — exposure predicts agreement

TROBRIAND ISLANDS · PAPUA NEW GUINEA

- The wide-eyed gasping configuration — the proposed universal fear face — is read as a threat display, an intent to attack
- The scowl was most often described as wanting to avoid a social interaction
- In free labelling, none of the proposed configurations drew the expected label above chance
- Faces were labelled consistently — just with different words than predicted

Participants frequently described actions rather than inner states — a gasping face is someone looking. In several of these communities another person's mind is not assumed to be readable at all, a stance anthropologists call opacity of mind.

Barrett et al. (2019); Gendron et al. (2014, 2018); Crivelli, Jarillo & Fridlund (2016, 2017); Crivelli, Russell et al. (2016).

THE VERDICT

Reliability and specificity: the paper's own summary

EXPRESSION PRODUCTION

	RELIABILITY	SPECIFICITY
Adults, developed, spontaneous, lab	weak	unknown
Adults, developed, spontaneous, natural	weak	unknown
Adults, developed, posed	weak–strong	unknown
Adults, remote, spontaneous	unclear	unknown
Adults, remote, posed	weak–strong	unknown
Newborns, infants, toddlers	unsupported	unsupported
Congenitally blind	unsup–weak	unsupported

EMOTION PERCEPTION

	RELIABILITY	SPECIFICITY
Developed, choice-from-array	mod–strong	unknown
Developed, reverse correlation	moderate	moderate
Developed, free labelling	weak–mod	weak
Developed, virtual humans	unknown	unknown
Remote, choice-from-array (pre-2008)	mod–strong	unknown
Remote, choice-from-array (post-2008)	weak–mod	unsupported
Remote, free labelling (pre-2008)	weak–strong	variable
Remote, free labelling (post-2008)	unsupported	unsupported
Infants, young children	unsupported	unsupported

Grey = not reported. Specificity is unknown or unsupported in every row but one.

Barrett et al. (2019), Table 4. Bands follow Haidt & Keltner (1999).

DISCUSSION 3

~6 minutes

Does the paper prove too much?

- The thresholds come from Haidt and Keltner — sympathetic to the basic emotion tradition. Are they the right bar for an applied system, or too demanding? Would we hold any behavioural signal to them?
- Population-level reliability is weak. Does that rule out person-specific or context-conditioned models — or is the paper's target only the context-free, one-size-fits-all inference?
- Much of the specificity column is unknown rather than refuted. How much should absence of evidence move us?

5

Implications

What this means for the systems we build

GUARD RAILS

What the paper is not claiming

Not: facial movements are random	The configurations do occur above chance. There is signal. It is weaker, less specific, and more context-bound than assumed.
Not: faces are uninformative	Faces carry rich social information — intent, social motive, threat, affiliation. Trobriand participants read the gasping face consistently, just not as fear.
Not: emotions do not exist	The argument concerns the mapping between facial movement and emotion category, not the reality of emotional life.
Not: computer vision on faces is futile	The authors call these technologies powerful instruments for studying expression — deployed as measurement, not as mind-reading.
Not: nothing can ever be inferred	Facial movement measured in high-dimensional, dynamic context may yet serve the diagnostic purpose consumers want. That research has not been done.

Barrett et al. (2019), "Evaluation of the Empirical Evidence" and recommendations.

For our field

Reading out an internal state from facial movements alone, without context, is at best incomplete and at worst invalid — however sophisticated the algorithm.

The authors' framing: companies may be asking the wrong question. More data and better architectures do not repair an inference that is unlicensed in principle.

Scaling does not help

If the signal is context-dependent and the context is unmeasured, more images of faces without context add resolution to the wrong variable.

Benchmarks flatter us

Posed, exaggerated, forced-choice benchmarks are the conditions under which the common view looks strongest.

The label is the artefact

A model trained on annotator judgements learns the annotators' reverse inference, including its cultural specificity.

SCOPE

What survives the critique intact

The authors carve out applications that never depended on inferring an internal emotional state:

Clinical pain detection

The target is a behavioural signal with clinical utility, not a named emotion category.

Driver drowsiness

Eyelid and head dynamics predict a state with direct behavioural consequences.

Avatar and animoji driving

Reproducing facial movement is the entire goal — no inference is being made.

Face-based authentication

Identity, not state.

The dividing line: are you measuring a movement, or claiming to have read a mind?

Barrett et al. (2019), note 41.

CONSTRUCTIVE HALF

The research programme the authors propose

1	**Sample people deeply, not populations broadly**	Track individuals across many situations, times of day, and social settings to learn personal expressive repertoires — a within-person, big-data approach.
2	**Measure in high dimension**	Facial movement plus posture, gait, voice, ambulatory autonomic signals, eye tracking, wearable neuroimaging, and the physical and social properties of the space.
3	**Quantify context explicitly**	Manipulate or at least measure context. Theories about context cannot be adjudicated until context is measured and quantified.
4	**Let people annotate their own episodes**	Self-annotation on affective dimensions such as valence and arousal, and on appraisals, rather than forced choice among six category labels.
5	**Discover categories rather than verify them**	Use unlabelled and unsupervised approaches on large cross-cultural image sets instead of asking whether other cultures resemble the United States.
6	**Use virtual humans as instruments**	Contingent, controllable social partners give richer context than static photographs without abandoning experimental control.
7	**Report reliability and specificity, and use signal detection**	Formal detection analytics and information-theoretic measures rather than raw agreement rates; Bayesian methods so the null can be tested directly.

Barrett et al. (2019), "Recommendations for Future Scientific Research" and Table 9.

READING CRITICALLY

Pressure points in the paper itself

Contested reading of the evidence	Basic emotion researchers dispute the interpretation, arguing that large-scale naturalistic and multimodal studies reveal more structure than this review credits — and that emotion is expressed through more than the face.
Inherited thresholds	The reliability bands are borrowed from Haidt and Keltner and applied across very different designs. Whether a 70–90% bar is the right standard, and whether it is applied evenly, is arguable.
A review, not a meta-analysis	The synthesis is largely qualitative. Study quality and sample size are weighted by judgement rather than by a stated protocol.
Unknown treated as damning	Much of the specificity column reflects non-reporting. That is a real indictment of the literature, but it is not the same as a demonstrated failure.
Theoretical commitments	The lead author is the principal architect of the constructionist alternative. That is not disqualifying — but the target is also her opponents' position.
Is diagnosticity the right bar?	Few behavioural measures meet a fingerprint standard. If we required it of gait, voice, or language, would they survive?

Seminar commentary. PSPI target articles are published with peer commentary; see the accompanying responses in the same issue.

CLOSING

Five things to carry out of this room

1. **Distinguish the two inferences**

   p(face | emotion) is what the literature measures. p(emotion | face) is what your system claims. They are not interchangeable.

2. **Ask for the false positive rate**

   Reliability without specificity cannot support a reverse inference. If a paper or a vendor does not report it, that is the finding.

3. **Interrogate the label, not just the model**

   Ask where ground truth came from. If it came from annotators, say that your system predicts annotator judgement.

4. **Treat context as a variable, not noise**

   The unexplained variance is not error. It is the situation, the body, the relationship, and the culture — all measurable.

5. **Name what you measure**

   Facial configuration, not emotional expression. The honest name usually reveals what the system can and cannot support.

CLOSING DISCUSSION

~8 minutes

Where does this leave affective computing?

- If we accept the critique, what is the most defensible thing an emotion-aware system can claim? Is there a version of this technology you would still build?
- The authors want deep within-person, high-dimensional, context-rich data. What would it take to collect that ethically — and who would consent to it?
- Several jurisdictions now restrict emotion recognition in workplaces and schools. Is regulation the right instrument here, or is the problem better handled by measurement standards and honest labelling?
- If categories should be discovered rather than assumed: what would an unsupervised study of facial movement in the wild look like, and how would you evaluate it?

SOURCES

The paper and its surroundings

THE TARGET ARTICLE

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. Psychological Science in the Public Interest, 20(1), 1–68.

CITED WITHIN, WORTH READING DIRECTLY

Duran & Fernández-Dols (2018) on production meta-analysis · Elfenbein & Ambady (2002) on perception meta-analysis · Crivelli, Carrera & Fernández-Dols (2015) on judo winners · Gendron et al. (2014, 2018) on Himba and Hadza · Crivelli, Jarillo & Fridlund (2016, 2017) on Trobriand Islanders · Srinivasan & Martinez (2018) on many-to-many labelling · Siegel et al. (2018) on autonomic specificity

MEASUREMENT AND COMPUTER VISION

Ekman & Friesen (1978), FACS · Benitez-Quiroz et al. (2016), EmotioNet · Benitez-Quiroz et al. (2017), the EmotioNet Challenge · Martinez (2017) on automated facial action detection

THE OTHER SIDE OF THE ARGUMENT

Haidt & Keltner (1999) on the criteria adopted here · Ekman & Cordaro (2011) and Keltner & Cordaro (2017) on the basic emotion position · the peer commentaries published alongside this article in the same PSPI issue