Supplemental Materials

To accompany

Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures

Contents

I. Justification of linguistic exclusions

Initial analyses of linguistic abstractness scores including all participants revealed that the thin plate smoothing spline model, R^2 = .21, AIC = -167.63, provided a better fit than the null, AIC = -130.09, linear, β = .34, p < .001, R^2 = .11, AIC = -152.44, quadratic, β = .14, p = .034, R^2 = .13, AIC = -155.03, and cubic, β = -.20, p = .002, R^2 = .17, AIC = -162.59, models (**Figure S1**, log likelihood test comparing cubic and spline models: p = .012). However, further analyses suggested that linguistic abstractness scores for very young participants might not be valid. The relationship between linguistic abstractness scores and coder abstractness scores varied across age: there was a significant interaction between age and coder abstractness scores in predicting linguistic abstractness scores, β = .45, p < .001. Thus the concordance between coder and linguistic measures of abstractness were not equivalent across age when all participants were included. Specifically, concordance was weaker at younger ages and stronger at older ages.

We hypothesized that this breakdown in concordance at young ages may have arisen because there was little usable linguistic data on which linguistic abstractness measures could be computed for young children. Young participants both (i) spoke fewer words on average than older participants (see **Results**) and (ii) they had fewer usable trials to include in linguistic analyses (given that they were less likely to fully comprehend emotions and abstraction scores were only valid for emotions that participants fully comprehended). To test whether a paucity of linguistic data in young children could explain low concordance between human and linguistic abstractness scores at young ages, we computed the total word count for all usable emotion definitions for each participant. Indeed, total usable word count increased across age, β = .28, p < .001. However, controlling for usable word count reduced the interaction between age and coder abstractness scores in predicting linguistic abstractness scores to a statistical trend, β = .23,

p = .052. Hence, accounting for age-related differences in usable word count helped stabilize concordance between coder and linguistic measures of abstractness across age.

Consequently, we excluded 7 participants aged 4.13-6.91 whose usable word count was < 100 usable words (3.57% of participants). Following this exclusion, the concordance between coder and linguistic measures of abstractness was equivalent across age: there was no longer a significant interaction between age and linguistic abstractness scores in predicting coder abstractness scores, β = .19, p = .157. Researchers should note that having sufficient linguistic data appears to be important for computing stable linguistic abstractness scores.

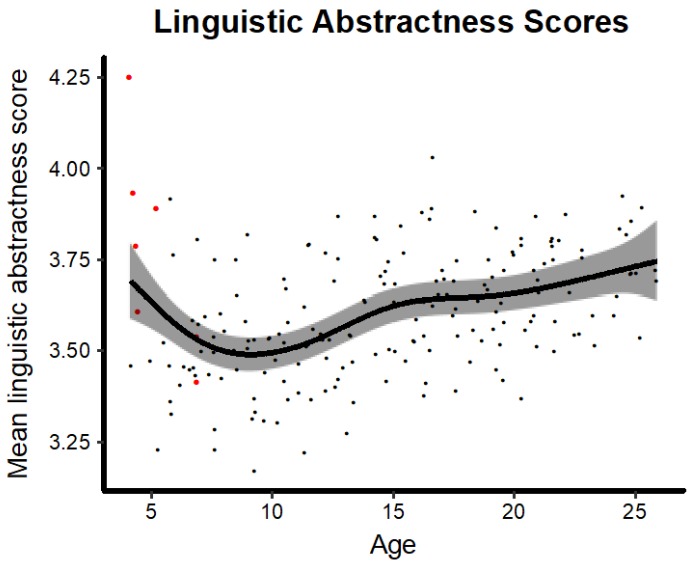

Figure S1. Linguistic abstractness scores across age when including data from all participants. The 7 red points represent participants who were excluded from primary analyses due to having overall low levels of usable linguistic data. Excluding these participants stabilized the correlation between coder abstractness scores and linguistic abstractness scores across age. Line of best fit and 95% CIs (grey shading) represent results from the thin plate smoothing spline model.

II. Justification of plateau definition

Because the field has not developed standard thresholds for defining plateaus of nonlinear fit lines, we created a definition of a plateau by adapting mathematical principles for use with empirical data. As stated in the main text, a plateau can be mathematically defined as occurring when (i) there is reduced rate of change in the dependent variable and (ii) the dependent variable reaches the maximum/minimum of its development. However, it was important that this mathematical definition could be flexibly applied across datasets, especially given that we applied it to bootstrapped simulated datasets. Thus we began with "mathematical ideals" and ended our search when we had discovered a definition that could accurately detect plateaus across datasets.

An ideal mathematical definition of a plateau would be the point at which the first derivative of the curve reaches 0. However, applying this definition to empirical data—which are far noisier than mathematical ideals—led to clearly inaccurate definitions of the plateau in some bootstrapped simulations. One such simulated bootstrap sample is presented in **Figure S2** below. As you can see, the slope of the curve clearly slows and flattens in childhood, but the 0%/year threshold (red line) is only reached in young adulthood. Thus this threshold can be set too strictly to produce an accurate definition. As further evidence that this threshold is too strict, around 7% of bootstrapped simulations of example situation use supposedly did not reach a plateau within the measured age window, but an examination of these simulated samples suggested otherwise: the curve clearly plateaued, but it did not reach the 0%/year threshold.

Consequently, we adjusted this threshold to reasonable values (e.g., 0.05%/year, 0.10%/year, 0.50%/year, 1%/year) and tested how well these thresholds fit the data. We found that the 1%/year threshold was slightly too lenient, as the plateau was identified before the slope

appeared to fully flatten (see blue line in **Figure S2**). Thus we selected 0.5%/year as the threshold because it provided a good balance between these two demands (see black line **Figure S2**). This not only removed inaccuracies in the emotion comprehension score bootstraps (as depicted in **Figure S2**), it also reduced "unidentified" plateaus in bootstrap simulations from 5.84% to 0.84% for example situation use, from 1.26% to 0% for coder abstraction scores, and from 0.11% to 0% for general definition use.

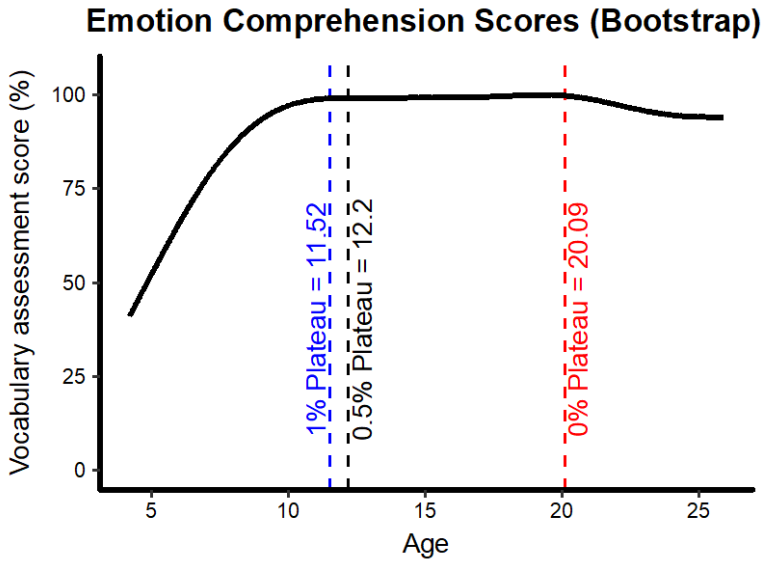

Fig S2. One bootstrapped sample of emotion comprehension scores (i.e., a randomly drawn sample from the study's data with replacement that was used to make bootstrapped confidence intervals of plateaus) that illustrates the influence of the first-derivative threshold on how plateaus are identified. Setting the threshold to 0%/year causes the plateau to be identified almost 8 years after the curve appears to flatten (red dashed line), and setting the threshold to 1%/year causes the plateau to be identified slightly before the curve appears to flatten (blue dashed line). Hence, 0.5%/year (black dashed line) was chosen to provide a well-fitting balance between these demands.

However, this criterion alone was not sufficient for accurately defining plateaus. In particular cases, the dependent variable had a slope that showed only a *temporary* flattening far before the variable hit a final plateau point (e.g., at age 6.76, see red line in **Figure S3** below). Even though change in the dependent variable slowed to < 0.5%/year at this point, scores increased again after this temporary flattening and then eventually plateaued at the end of

adolescence. Thus we added a second criterion such that plateaus were only characterized once the age-related change had completed. We defined this as within the 95% CI of the smoothing spline's estimate of the maximum value of the curve (i.e., within the grey window surrounding the highest point of the black line). This ensured that a "plateau" wasn't when the curve flattened *momentarily* but when it flattened at the peak of development. We observed that adding this second definition ensured that all plateaus were defined accurately (black dashed line in **Figure S3**).

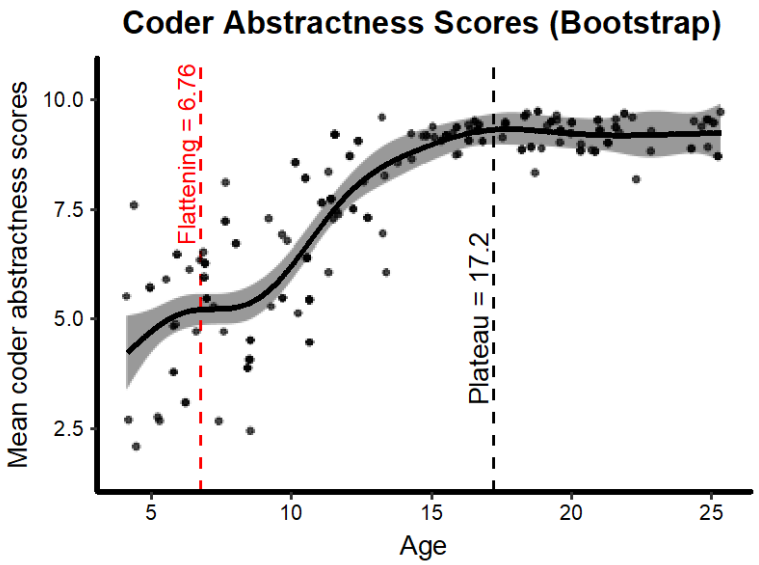

Figure S3. One bootstrapped sample of coder abstractness scores (i.e., a randomly drawn sample from the study's data with replacement that was used to make bootstrapped confidence intervals of plateaus) that illustrates the influence of the additional criterion for identifying plateaus (the value of the spline regression must be within the 95% CI of the maximal value of the spline) on how plateaus are identified. Without this criterion, any temporary flattening of the spline curve could be inaccurately labeled a plateau (red dashed line), but with this criterion, the accurate plateau is identified (black dashed line).

This is the process by which we created a specific statistical definition from the general mathematical principles of plateaus. Because the bootstrapping approach forced us to test the stability of this definition across 10,000 samples of 4 different dependent variables, there is reason to believe it can locate plateaus with high accuracy, stability, and generalizability.

III. Analyses only including participants who fully comprehended all emotion words

Measures of emotion abstraction presented in the manuscript only include data from emotions that participants fully comprehended. Because younger participants comprehended fewer emotion terms than older participants (see **Results**), abstraction scores may reflect different sets of emotions for different ages. This could result in lower abstraction measures in younger ages because the emotions they understood were less abstract. Although we conceptualized this possibility as a part of the phenomenon this study investigates, we conducted supplemental analyses to ensure that age-related increases in emotion abstraction were not *solely* due to age-related differences in emotion comprehension. We used linear regressions to test whether each dependent variable was significantly related to age when restricting the sample to participants who (i) fully comprehended all 24 emotions and (ii) were within the active developmental window for that dependent variable.

For dependent variables that followed emergent developmental patterns (i.e., coder abstractness scores, general definition use, and example situation use), we restricted the sample to the 64 participants aged 6.24-17.54 who fully comprehended all 24 emotion words and were younger than the plateau of coder abstractness scores (M = 12.85, SD = 2.98, 53.13% female). We used an age ceiling to restrict this sample to the active window of increasing emotion abstraction. Linear regressions revealed that age was significantly related to coder abstractness scores, β = .79, p < .001, general definition use, β = .72, p < .001, and example situation use, β = -.68, p < .001, even within this restricted sample.

For dependent variables that followed linear developmental patterns (i.e., linguistic abstractness scores and synonym use), we restricted the sample to the 113 participants aged 6.24-25.91 who fully comprehended all 24 emotion words (M = 16.36, SD = 4.81, 53.10% female).

No age ceiling was placed on this subset of participants, as these dependent variables increased linearly across the full sample (see **Results**). Again, linear regressions revealed that both linguistic abstractness scores, $\beta = .45$, $p < .001$, and synonym use, $\beta = .26$, $p = .006$, significantly increased across age even within this restricted sample.

These supplementary analyses suggest that age-related differences in emotion comprehension do not confound analyses of emotion abstraction presented in the main text. Age-related changes in emotion abstraction arose even when emotion comprehension was held constant.

IV. Analyses excluding trials on which participants only provided synonyms

We originally counted synonyms as demonstrating full comprehension because i) Wechsler vocabulary assessments (on which this test was modeled) grant synonyms full credit for several items (Wechsler, 1999) and ii) providing synonyms demonstrates at least some understanding of how words within a taxonomy should be organized (i.e., that two words refer to the same concept and not different concepts). However, only providing a synonym might reflect shallow semantic priming rather than full comprehension. Hence, we conducted control analyses to protect against the possibility that synonym use did not reflect full comprehension.

Overall, the percentage of trials on which participants only provided synonyms was 4.49% (4.90% of 2-point responses). We analyzed relations between age and the tendency to only provide synonyms (quantified as the number of trials on which participants only provided a synonym) as we did for all other variables. AICs suggested that the smoothing-spline model provided a slightly better fit than the null model (**Table S1**). However, three points of evidence suggest that there is not significant variation in the tendency to only provide synonyms across age: i) there was no significant linear relationship with age (p = .193), ii) the R^2 of this model was only .01, and iii) no plateau could be defined because the rate of change of this variable was always lower than the plateau threshold of .5%/year (i.e., according to the plateau definition set for other variables, this variable doesn't change at all over the age window we sampled).

Table S1. Results of null, polynomial, and thin plate smoothing spline analyses of the tendency to only provide synonyms.

Dependent variable	Model	β	p	Adj R^2	AIC
Tendency to only provide synonyms (% of trials)	Null				1327.46
	Linear	.09	.193	.004	1327.74
	Quadratic	-.08	.243	.006	1328.35
	Cubic	.10	.163	.01	1328.36
	Spline			**.01**	**1327.28**

Note: Bold text indicates best fitting model for each dependent variable, as determined by having the lowest AIC. β = standardized beta, Adj R^2 = adjusted R^2, AIC = Akaike Information Criterion

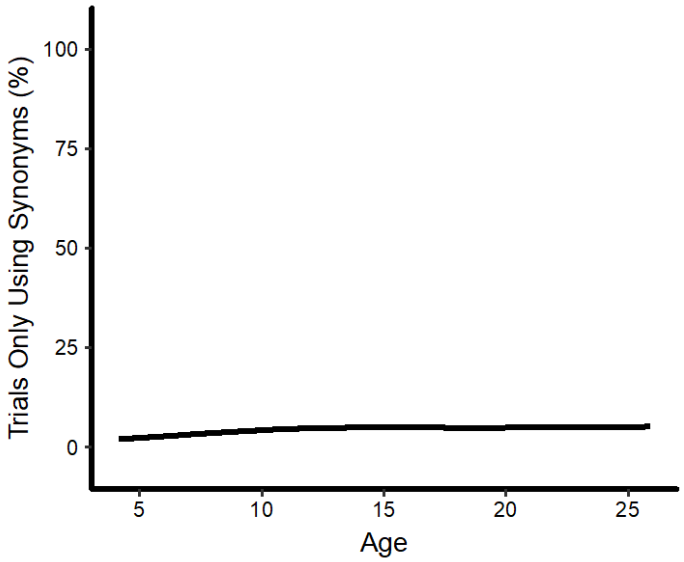

Figure S4. The percentage of trials on which participants only provided synonyms presented across age. The black line and grey shaded region present the fit of the smoothing spline regression and 95% CI, respectively.

We then conducted a set of analyses to test how dramatically results would shift if we assumed that trials on which participants only provided synonyms actually did not reflect comprehension the emotion words. We recoded all trials on which participants only provided synonyms as 0-point responses (which necessitated removing all abstraction measures from these trials) and recomputed all dependent variables. Overall, we find that patterns of results are very similar to what is presented in the main text, and all major conclusions remain unchanged.

The best fitting age models do not change for all dependent variables (note that synonym use was not analyzed as a dependent variable for these analyses; **Table S2**). Emotion comprehension scores, coder abstractness scores, general definition use, and example situation use followed emergent patterns; linguistic abstractness scores increased linearly across development; and physiological marker use was elevated in childhood. Although rescoring synonym-only responses as incorrect shifted the plateau of emotion comprehension scores slightly earlier (from age 10.95 to age 10.17), abstraction plateaus remained stable: the coder

abstractness score plateau remained at age 17.69, general definition use plateau shifted from age 17.19 to age 17.08, and the example situation use plateau remained at age 18.19.

Additionally, the significance of bootstrapped comparisons between plateaus remained identical. The emotion comprehension plateau remained significantly earlier than the plateaus of coder abstractness scores, 95% CI = [3.69, 11.00], general definition use, 95% CI = [3.90, 11.52], and example situation use, 95% CI = [4.83, 15.00]. Additionally, bootstrapped comparisons suggested no significant differences between plateaus for measures of emotion abstraction: coder abstractness scores vs. general definition use, 95% CI = [-1.79, 3.56], coder abstractness scores vs. example situation use, 95% CI = [-4.43, 0.59], and general definition use vs. example situation use, 95% CI = [-1.25, 5.48]. Note that these bootstrap comparisons were calculated after removing 0.65% of example situation simulations in which a plateau was not detected.

These results suggest that conclusions presented in the main text are not affected by the possibility that providing a synonym might not reflect full emotion word comprehension. Nonetheless, it may be wise for future researchers who use the emotion vocabulary assessment to only grant 2 points to responses that include example situations or general definitions.

Table S2. Results of null, polynomial, and thin plate smoothing spline models for each dependent variable after removing trials on which participants only provided synonyms.

Dependent variable	Model	β	p	Adj R^2	AIC
Emotion comprehension scores	Null				1631.82
	Linear	.41	< .001***	.16	1597.78
	Quadratic	-.43	< .001***	.35	1550.54
	Cubic	.26	< .001***	.41	1530.98
	Spline			**.45**	**1519.10**
Coder abstractness scores	Null				847.61
	Linear	.77	< .001***	.59	672.65
	Quadratic	-.38	< .001***	.73	590.42
	Cubic	-.01	.785	.73	592.34
	Spline			**.75**	**583.25**
Linguistic abstractness scores	Null				-133.81
	Linear	**.48**	**< .001*** **	**.23**	**-180.81**
	Quadratic	-.0005	.994	.22	-178.81
	Cubic	-.07	.278	.22	-178.01
	Spline			.23	-180.81+
General definition use	Null				1860.86
	Linear	.73	< .001***	.54	1710.86
	Quadratic	-.44	< .001***	.73	1608.55
	Cubic	.09	.019*	.73	1604.90
	Spline			**.74**	**1601.39**
Example situation use	Null				1881.05
	Linear	-.69	< .001***	.47	1758.78
	Quadratic	.33	< .001***	.57	1716.12
	Cubic	.02	.642	.57	1717.90
	Spline			**.58**	**1713.33**
Physiological marker use	Null				1403.99
	Linear	-.24	< .001***	.05	1394.74
	Quadratic	-.006	.933	.05	1396.73
	Cubic	.12	.084	.06	1395.67
	Spline			**.09**	**1390.08**

Note: Bold text indicates best fitting model for each dependent variable, as determined by having the lowest AIC. + In these two cases, spline model algorithms revealed that the best fits were identical to linear models; hence we interpreted the linear models. β = standardized beta, Adj R^2 = adjusted R^2, AIC = Akaike Information Criterion, *** $p < .001$, ** $p < .01$, * $p < .05$

V. Emotion vocabulary assessment

Materials. 27 laminated index cards with emotions written on them, scoring manual, scoring sheet, and computer for audio recording. Emotions: amazed, angry, annoyed, bored, calm, disappointed, disgusted, embarrassed, excited, grumpy, happy, hate, jealous, lonely, love, nervous, pleased, proud, relaxed, sad, safe, scared, sorry, surprised, thankful, upset, and worried.

General recommendations. Make sure that the deck of emotion cards has been shuffled multiple times prior to starting the task so that emotion order is random. Avoid providing corrective or praising feedback in this task. Attempt to keep them motivated but not bias their answers. Use encouraging but neutral feedback. For example, you might say things like "ok", "let's try another", or "you're doing good work" rather than "that's correct." Try to only look at the scoring sheet while they are speaking (i.e., don't give them feedback via eye contact), and try to resist nodding or giving other forms of feedback. Try not to cut them off after they provide a suitable response, as they should share whatever they think about each emotion.

Procedure. Begin by providing them with the following instructions: "This game is called the Word Definitions Game. In this game, I'm going to say a word, and I want you to do your best to tell me what the word means. I won't be able to talk to you about the game until we're all done with the study. Also, we should only spend a few minutes on this game, so try to give each definition in about 30 seconds or less. Are you ready to start?" Start the audio recorder. On each trial, pull a card from the deck of emotion terms, show it to them, and say: "What does _____ mean?" Follow the scoring guidelines below for each trial and probe them if their responses do not initially earn 2 points.

Overall scoring guidelines. Two-point answers include: (i) a reasonable description of the definition, even if the definition is not *exact* (see example definitions below), (ii) a synonym

(even if colloquial) or closely related emotional term (e.g., “pissed off” or “mad” for angry), or (iii) an example situation that would reasonably evoke the emotion in question and not other emotions (e.g., “I felt sad when my pet died” for sad). If the participant provides an example situation that could evoke multiple emotions (e.g., “I felt angry when my sister got something that I wanted,” which could evoke jealousy or anger), query them to ask for a more specific definition or other situations that can confirm that they know what the word means. **Note:** It is possible that synonym use actually represents shallow semantic priming rather than true comprehension of an emotion word’s underlying concept. Although we found that trials on which participants only provided synonyms were rare (4.49% of all trials), we encourage future researchers to consider not counting synonyms as 2-point responses to avoid the possibility of mistakenly accepting priming as comprehension.

One-point answers include a response that is of the correct valence but overly vague. No features are provided to distinguish this emotion from other similar emotions (e.g., “good,” “bad,” “positive,” “negative,” etc.). This includes responses that involve situations that could give rise to several emotions of a similar valence as the target emotion. If participants’ responses only earn 1 point, respond with a query (e.g., “can you tell me more about [emotion]?” or “can you tell me more about that situation?”). Repeat queries as needed to probe participant’s understanding of the emotion term. Score the participant’s response based on their responses to queries (i.e., if a 2-point response is given in response to queries, they earn 2 points for this emotion, even if their initial response earned 1 point or 0 points).

Zero-points are awarded when it is clear that the participant does not know what the emotion word means, even after probing. These responses include: (i) an incorrect definition of the term, (ii) an example of a situation that is unlikely to evoke the emotion in question (e.g., “I

feel happy when my pet dies"), or (iii) the response "I don't know." For all zero-point answers, query to thoroughly assess the participant's understanding of the emotion term. Score their response based on their responses to queries (i.e., they can earn 1 or 2 points for the emotion if later responses clarify their understanding of the emotion term). Some example queries include saying (i) "it's ok if you don't know all of these words, but let's just make sure. What do you think [emotion] means?", (ii) "What are some other feelings that might be like [emotion]?", (iii) "What are some other words people use to describe feeling [emotion]?", (iv) "What are some things that might make someone feel [emotion]?", (v) "Have you ever heard of someone feeling [emotion] in a book or movie? What made them feel that way?" However, use these latter forms of encouragement (probes for synonyms/situations) sparingly. One goal of the assessment is to measure how much people generate these types of strategies on their own, so only use them if thorough probing suggests they are not able to spontaneously generate a full definition.

Definition and synonym guidelines. Definitions below are based on: Merriam-Webster Dictionary for Adults, Merriam-Webster Dictionary for Children, Oxford American Dictionary, and Google Dictionary. Synonyms are taken from a variety of sources. These definitions are only to serve as a rough guide in your scoring.

***Angry*.** Definition: strong feeling of being upset, annoyed, displeased, or hostile. Synonyms: irate, mad, annoyed, cross, vexed, irritated, indignant, irked, furious, enraged, infuriated, in a temper, displeasure, fury, aggravated, livid; ticked off, pissed off; losing one's temper.

***Annoyed*.** Definition: slightly angry; irritated. Synonyms: irritated, vexed, angry, exasperated, irked, get on someone's nerves, ruffle someone's feathers.

***Amazed*.** Definition: surprised greatly; confused; filled with wonder; filled with

astonishment. Synonyms: astonished, astounded, surprised, stun, stagger, shock, stupefy, awe, stop someone in their tracks, leave open-mouthed, leave aghast, take someone's breath away…

Bored. Definition: weary or restless because one is unoccupied or lacks interest in one's current activity. Synonyms: disinterested, dull, sick and tired, fatigued, blasé, spiritless, tired, inattentive, turned off.

Calm. Definition: in a quiet and peaceful state or condition; not feeling or showing nervousness, anger or other emotions. Synonyms: serene, tranquil, relaxed, unruffled, unperturbed, unflustered, untroubled…

Disappointed. Definition: sad, unhappy, or displeased because someone or something has failed to fulfill one's hopes or expectations. Synonyms: upset, saddened, let down, cast down, disheartened, downhearted, downcast, depressed, dispirited, discouraged, despondent, dismayed, distressed.

Disgusted. Definition: a strong feeling of dislike for something that has a very unpleasant appearance, taste, smell, etc.; annoyance and anger that you feel toward some behavior that is not good, fair, appropriate; revulsion or profound disapproval. Synonyms: revolt, repel, repulse, sicken, nauseate, turn someone's stomach.

Embarrassed. Definition: confused and foolish in front of other people; self-consciousness, shame, or awkwardness. Synonyms: mortified, red-faced, blushing, abashed, shamed, ashamed, humiliated, awkward, self-conscious, uncomfortable…

Excited. Definition: eager enthusiasm and interest. Synonyms: thrilled, exhilarated, animated, enlivened, electrified.

Grumpy. Definition: easily annoyed or angered; having a bad temper or complaining often; sulky. Synonyms: crabby, ill-tempered, short-tempered, crotchety, testy, irritable, prickly,

grouchy, snappy, cranky, irritable.

***Happy*.** Definition: pleasure and enjoyment because of life, situation, etc.; contentment. Synonyms: cheerful, cheery, merry, joyful, jovial, jolly, jocular, gleeful, delighted, untroubled, smiling, beaming, grinning, in good spirits, in a good mood, lighthearted, pleased, content, satisfied, gratified, sunny, joyous.

***Hate*.** Definition: intense or passionate dislike (for someone); have a strong aversion to (something). Synonyms: loathe, detest, despise, dislike, abhor, repelled by…

***Jealous*.** Definition: intolerant of rivalry or unfaithfulness; envy of someone or their achievements and advantages. Synonyms: envious, covetous, desirous.

***Lonely*.** Definition: sad from being apart from other people; sad because one has no friends or company; solitary. Synonyms: isolated, alone, lonesome, friendless, with no one to turn to, abandoned, rejected, unloved, unwanted, outcast, gloomy, sad, depressed, desolate, forsaken.

***Love*.** Definition: strong or constant affection for a person; feel a deep romantic or sexual attachment to someone. Synonyms: deep affection, fondness, tenderness, warmth, intimacy, attachment, endearment, care very much for, hold very dear, adore, think the world of, be devoted to, dote on, worship.

***Nervous*.** Definition: worried and afraid about what might happen; easily agitated or alarmed; tending to be anxious; highly strung. Synonyms: high-strung, anxious, edgy, tense, excitable, jumpy, skittish, brittle, neurotic.

***Pleased*.** Definition: happy or satisfied. Synonyms: happy, glad, delighted, grateful, thankful, content, satisfied.

***Proud*.** Definition: deep pleasure, satisfaction, or happiness as a result of one's own

achievements, qualities, or possessions or those of someone with whom one is closely associated; attitude or people who think that they are better or more important than others. Synonyms: pleased, glad, happy, delighted, joyful overjoyed, thrilled, satisfied, gratified, content.

Relaxed. Definition: calm and free from stress, worry, or anxiety; free from tension and anxiety; at ease. Synonyms: comfy, cozy, relaxed, content, satisfied, peaceful, resting, easygoing, undisturbed.

Sad. Definition: grief or unhappiness; sorrow. Synonyms: unhappy, sorrowful, dejected, depressed, downcast, miserable, down, blue, down in the dumps, blah.

Safe. Definition: free from harm or risk; secure form threat of danger, harm, or loss. Synonyms: secure, protected, shielded, sheltered, guarded, out of harm's way.

Scared. Definition: afraid of something; nervous, frightened, fearful. Synonyms: afraid, startled, nervous, fearful, panicky, alarmed, intimidated, terrified, petrified, terrorized, spooked.

Sorry. Definition: sorrow or regret; distress, especially through sympathy with someone else's misfortune. Synonyms: sad, unhappy, sorrowful, distressed, upset, downcast, downhearted, despondent.

Surprised. Definition: feelings caused by something that is unexpected or unusual. Synonyms: amaze, astonish, astound, dumbfounded, shocked, startled.

Thankful. Definition: glad that something has happened or not happened, that something or someone exists, etc. pleased and relieved; gratitude. Synonyms: grateful, appreciative, filled with gratitude, relieved.

Upset. Definition: unhappy, worried; disappointed. Synonyms: distress, trouble, unsettled, worried, bothered, agitated, hurt, sadden, grieve, perturbed, disquieted.

Worried. Definition: fear or concern because you think that something bad has happened or could happen; anxious, upset, or troubled about actual or potential problems. Synonyms: anxious, perturbed, troubles, bothered, concerned, upset, distressed, uneasy, agitated, nervous, edgy, tense, keyed up, jumpy, stressed, strung out.