

Teacher key

Condition column shows Form A / Form B. Tally results by condition: the comparisons that matter are open vs. first-step (B), the three stakes notes (C), choice vs. assigned (D), and answer-only vs. steps (G). Paper gives coarser data than the online version: no per-screen timing, and no record of how long a student sat before starting.

#	Item	Condition (A / B)	Answer	What each wrong choice means
Reading checks (not graded)				
1	R1		Purple	
Warm-up				
2	FL1		12	
3	FL2		6	
4	FL3		18	
5	FL4		17	
6	A1		14	
7	A2		60	
8	A3		30	
9	A4		72	
10	A5		10	
11	A6		21	
12	A7		3n	
13	A8		8	
14	A9		9	
Investigations				
15	T1	open / forced	77	
16	T2	forced / open	122	
17	T3	forced / open	152	
18	T4	open / forced	97	
Quick problems				
19	C1	low / credit	18	
20	C2	teacher / teacher	5	
21	C3	credit / low	27	
22	C4	credit / low	80	

#	Item	Condition (A / B)	Answer	What each wrong choice means
23	C5	teacher / teacher	10	
24	C6	low / credit	16	
Which deal is better?				
25	D1	choice / assigned	6	
26	D2	assigned / choice	8	
27	D3	assigned / choice	9	
28	D4	choice / assigned	12	
Reading checks (not graded)				
29	R2		Making posters	
Multi-step problems				
30	G1	trap / stepped	\$66	\$60 = stopped before the last step; \$68 = combined the percents; \$88 = used the wrong starting amount
31	G2	stepped / trap	33	18 = skipped a step; 30 = stopped before the last step; 363 = wrong operation
32	G3	stepped / trap	38	34 = wrong operation; 60 = stopped before the last step; 98 = stopped before the last step
33	G4	trap / stepped	9	15 = skipped a step; 18 = stopped before the last step; 6 = wrong operation
One longer problem				
34	K1		a) 260 b) 480 c) 220 d) 9 e) 80	
Reasoning				
35	H1		Dana is not in Ms. Ortega's class.	Dana is in Ms. Ortega's class. = reversed the logic; Dana did not study. = added a claim not given; The quiz was too hard. = added a claim not given
36	H2		Ben	Ana = mis-ordered; Cal = mis-ordered; Dee = mis-ordered
37	H3		E and 7	E and 4 = checked the confirming case; E only = checked only one case; All four cards = checked everything
38	H4		Students who eat breakfast may also differ in other ways, like sleep or home routines.	Ten points is too small to matter. = dismissed the claim; Test scores can't really be measured. = dismissed the claim; Breakfast is unhealthy for most students. = added a claim not given

#	Item	Condition (A / B)	Answer	What each wrong choice means
39	H5		Step 1	Step 3 = blamed a later step; Step 2 = blamed a later step; There is no mistake. = saw no mistake
40	H6		Ava tells the truth; Ben lies.	Ben tells the truth; Ava lies. = misjudged who tells the truth; Both are liars. = misjudged who tells the truth; Both are truth-tellers. = misjudged who tells the truth
Different subjects				
41	F1		12	
42	F2		47	
43	F3		She had planned for the dog's visit.	
44	F4		6	
Getting harder				
45	E1		42	
46	E2		41	
47	E3		66	
48	E4		Any valid expression equal to 100, e.g. $123 + 4 - 5 + 67 - 89$ or $123 - 45 - 67 + 89$	
Reading checks (not graded)				
49	R3		Skip ahead	

Reading checks: R1 correct = Purple (4th). R2 = Making posters (2nd). R3 = Skip ahead (4th). **Floor items** (FL) are deliberately trivial. A miss on a floor item with real effort means below the test's floor.

On paper: mark an investigation (B, D, F, E) as "started" if the work area or answer line has anything in it. Note which students tick "I used a nudge," and whether a student tried again on E1 to E3 after a wrong try (look at Try 1, 2, 3).